STIC-ILL

QP601. C72

Fr m: Lacourciere, Karen
Sent: Wednesday, July 24, 2002 10:31 AM
To: STIC-ILL
Subject: ill order

Please provide the following reference for use in examining 09/836,410

FASEB Journal, (March 12, 1999) Vol. 13, No. 4 PART 1, pp. A531.

IOVS, (March 15, 2001) Vol. 42, No. 4, pp. S94. print.

INVESTIGATIVE OPHTHALMOLOGY AND VISUAL SCIENCE, (2001 Nov) 42 (12) 3000-7.

Meth. Enzymol. 303, 19-44, (1999).

Genome Res. 10 (10) 1617-1630 (2000)

Genome res. 10 (11) 1757-1771 (2000)

Thank-You!

Karen A. Lacourciere Ph.D.
CM1 11D09 GAU 1635
(703) 308-7523

· state, or therapeutic efficacy. A
:red in single neurons affected by
le clues to pathogenetic disease
nterventions. The use of mRNA
ıpeutics is called transcript-aided
pled with single-cell resolution,
in diagnosis of disease states, as
:velop therapeutic strategies. For
idual diseased cell may increase,
is possible that a pharmaceutical
's may be specifically designed to
state. Alternatively, if functional
:hanges, then drugs targeting the
these altered mRNAs may prove
an approach is that information
d in a diseased cell may provide
drugs. For example, if the abun-
:eptor is altered as shown by the
ıdy available adrenergic receptor
ously been used in this particular
)eutically efficacious.
is a measure of the potential for
ly the functional entities within
:en result in disease. The ability to
pression, in single phenotypically
: will provide unique insight into
ınd will likely furnish unforeseen
anisms that produce disease eti-

ıtist **1,** 200 (1995).
rwine, *Dev. Genet.* **14,** 137 (1993).

[2] High-Efficiency Full-Length cDNA Cloning

By PIERO CARNINCI and YOSHIHIDE HAYASHIZAKI

Introduction

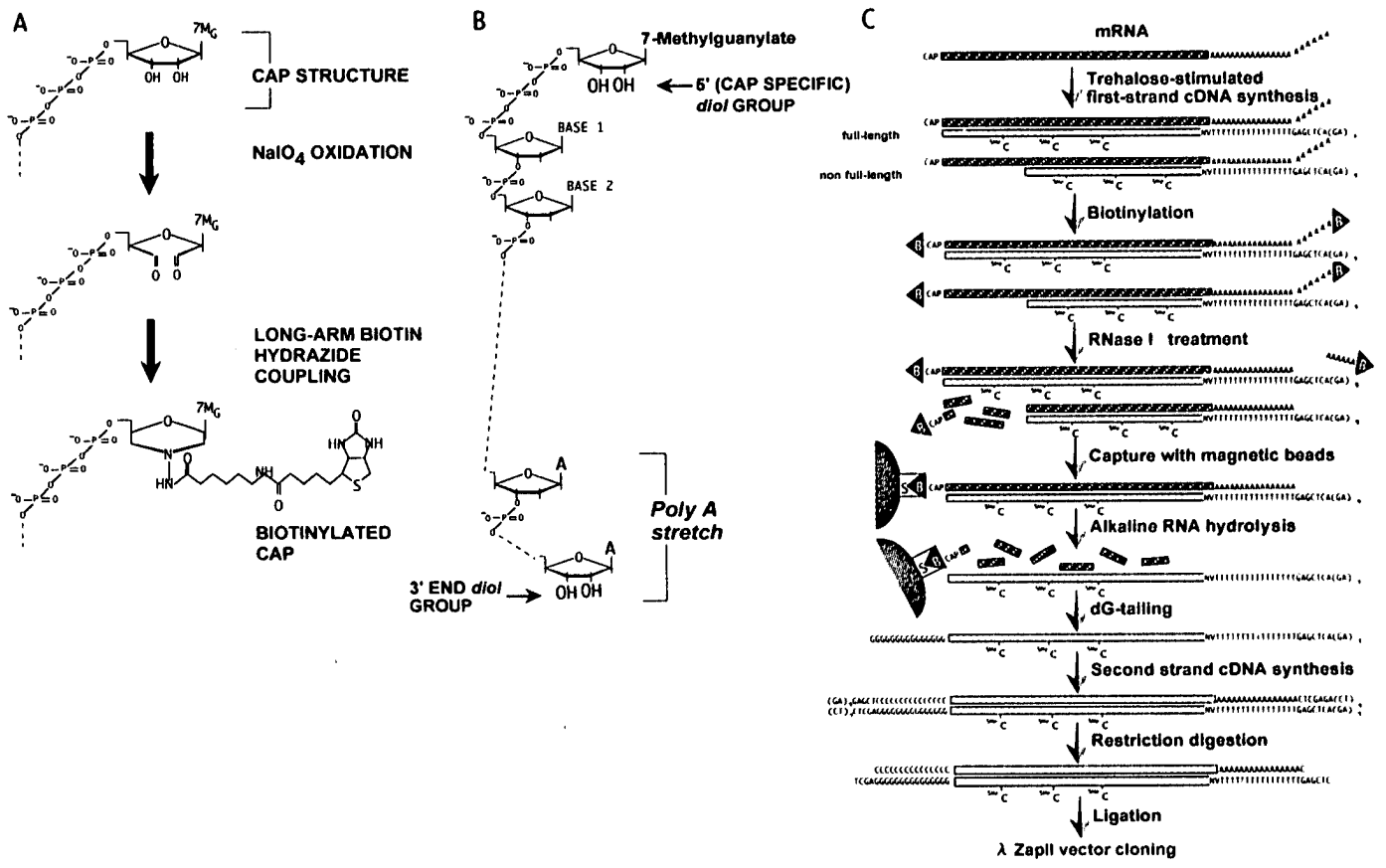
A full-length cDNA library is advantageous in that it allows cloning of a complete sequence in a single step. However, the representation of full-length cDNA clones has been low in cDNA libraries prepared using standard techniques. Full-length cDNA libraries have become available with a high proportion of the clones containing all of a particular coding sequence and its 3′ and 5′ untranslated regions (UTRs). Such libraries are particularly useful for large-scale sequencing projects (EST), in which the recovery of full-length clones from among truncated clones can be a formidable task.

In the preparation of full-length cDNA libraries, two problems have arisen. The first is the difficulty in reaching the cap site (5′ end of mRNA) in the first-strand synthesis with the reverse transcriptase (RT), the first enzyme involved in the preparation of a cDNA library. This is due to the presence of strong secondary structures of mRNAs.[1] These structures cause early termination of the reaction and detachment of the RT in a large number of clones, thus resulting in clones lacking the 5′ end, especially when priming the mRNA on the 3′ poly(A) tail. We have been able to overcome this drawback by introducing a disaccharide, trehalose, into the RT reaction. Trehalose can stabilize several enzymes, including RT, at temperatures as high as 60°[2] instead of the more usual 42° reaction temperature.[3] At 60°, the strength of the secondary structures of RNAs is thought to be greatly decreased, thus allowing much more efficient reverse transcription even of palindromic sequences, which are often encountered in the 5′ UTR of mRNAs. This results in longer cDNAs, higher representation of long, full-length cDNAs in the library, and an overall higher yield of the recovered full-length cDNA.

The other problem in library preparation is that there have not been effective methods for selection of full-length cDNAs from incompletely extended cDNAs. To solve this, we introduced a modified "biotinylated cap trapper" to select full-length cDNAs after biotinylation of the cap

[1] F. Payvar and R. T. Schimke, *J. Biol. Chem.* **254,** 7636 (1979).
[2] P. Carninci, Y. Nishiyama, A. Westover, M. Itoh, S. Nagaoka, N. Sasaki, Y. Okazaki, M. Muramatsu, and Y. Hayashizaki, *Proc. Natl. Acad. Sci. U.S.A.* **95,** 520 (1998).
[3] M. S. Krug and S. L. Berger, *Methods Enzymol.* **152,** 316 (1987).

0076-6879/99 $30.00


A
CAP STRUCTURE
NaIO4 OXIDATION
LONG-ARM BIOTIN HYDRAZIDE COUPLING
BIOTINYLATED CAP
B
7-Methylguanylate
5' (CAP SPECIFIC) diol GROUP
BASE 1
BASE 2
Poly A stretch
3' END diol GROUP
C
mRNA
Trehalose-stimulated first-strand cDNA synthesis
full-length
non full-length
Biotinylation
RNase I treatment
Capture with magnetic beads
Alkaline RNA hydrolysis
dG-tailing
Second strand cDNA synthesis
Restriction digestion
Ligation
λ ZapII vector cloning

structure (Fig. 1).[4,5] The method is based on a two-step chemical biotinylation of the two diol groups (Fig. 1A), one on the cap structure, specific to eukaryotic mRNAs, and the other on the terminal 3′ end of any RNA (Fig. 1B). The mRNA is biotinylated after the first-strand cDNA synthesis.[5] In the case of incomplete first-strand synthesis, there is a single-stranded RNA (ssRNA) that connects the biotinylated cap and the truncated cDNA–mRNA hybrid (Fig. 1C). This ssRNA can be digested by RNase I, an ssRNA-specific ribonuclease, in the case of truncated cDNAs, resulting in physical separation of the biotinylated cap from the cDNA. When the cDNA extends to the cap structure, RNase I cannot remove the cap from the hybrid, allowing its capture by steptavidin-coated magnetic beads (Fig. 1C). The second biotin group at the 3′ end is always removed by RNase I. In fact, the poly(A) tail is usually much longer than the part protected by the 16 Ts of the [5′(GA)$_8$ACTCGAG(T)$_{16}$VN-3′] oligonucleotide used to prime the first-strand cDNA. The degenerate V and N bases, where V is G, A, or C, and N is any base, are used for exact priming at the beginning of the poly(A), just after the cleavage and polyadenylation site. This priming thus allows for effective degradation of the unprotected single-stranded 3′ end of the poly(A) by RNase I (Fig. 1C), which is usually much longer than the fraction covered by the 16 Ts of the primer, and consequent removal of the biotin group at the 3′ end. The long sequence of repeated GAs at the 5′ end of the oligonucleotides (both first- and second-strand primers) is designed to provide a reasonably long DNA segment upstream of the restriction site to overcome the limitations encountered in the restriction cleavage close to the DNA ends.

The full-length cDNA is subsequently released from the beads by alkali treatment, followed by oligo(dG) tailing and second-strand synthesis mediated by a long-range thermostable DNA polymerase mixture,[6] which makes

[4] P. Carninci, C. Kvam, A. Kitamura, T. Ohsumi, Y. Okazaki, M. Itoh, M. Kamiya, K. Shibata, N. Sasaki, M. Izawa, M. Muramatsu, Y. Hayashizaki, and C. Schneider, *Genomics* **37,** 327 (1996).

[5] P. Carninci, A. Westover, Y. Nishiyama, T. Ohsumi, M. Itoh, S. Nagaoka, N. Sasaki, Y. Okazaki, M. Muramatsu, C. Schneider, and Y. Hayashizaki, *DNA Res.* **4,** 61 (1997).

[6] S. Cheng, C. Fockler, W. M. Barnes, and R. Higuchi, *Proc. Natl. Acad. Sci. U.S.A.* **91,** 5695 (1994).

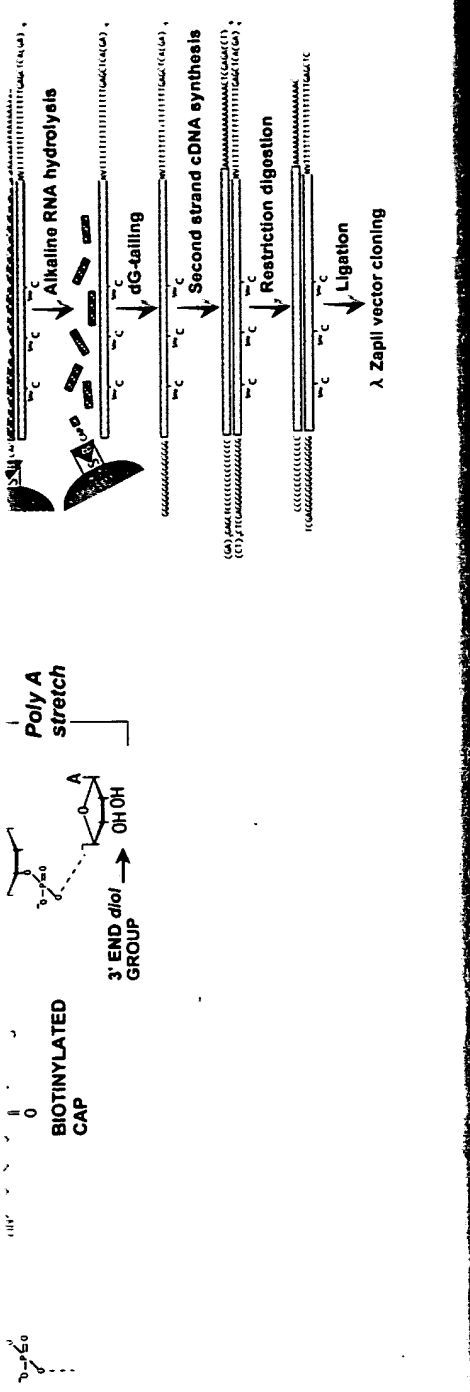


FIG. 1. Overall strategy for preparation of a full-length cDNA library by modified biotinylated cap trapper. (A) Two-step coupling of the biotin hydrazide with diol groups. (B) Structure of mRNA and position of the two diol groups at the 5′ end(cap) and the 3′ end of the mRNA. (C) Strategy for the preparation of the cDNA library. FL, Full-length first-strand cDNA; NFL, non-full-length cDNA.

it possible to synthesize the second strand with high efficiency even in the case of long cDNAs. Because this protocol employs 5-methyl-dCTP instead of dCTP for the first-strand cDNA synthesis, the resulting cDNA is hemimethylated and is resistant to restriction digestion by the enzymes *Xho*I[7] and *Sst*I.[8] These enzymes can cleave the unmethylated sites only on the first- and second-strand primer adapters. The cDNAs is then cloned in λ vector at high efficiency. λ insertional vectors are used here because of their high efficiency and high capacity (0–10 kb).[9] In addition, λ ZAPII is useful for automatic *in vivo* excision into pBluescript plasmid, which does not require tedious subcloning from λ DNA.[10,11]

In this chapter we describe protocols using the powerful "thermostabilized" RT, which make it possible to prepare efficiently full-length cDNAs longer than 10 kb, combined with the selection of cDNA by biotinylated cap trapper to remove residual non-full-length cDNAs. Using these protocols, we can prepare full-length cDNA libraries at high yield (several to tens of millions of independent clones can be routinely produced) without using polymerase chain reaction (PCR), which introduces sequence bias and causes overrepresentation of short clones in a library. For a more extensive discussion about alternative procotols for full-length libraries, see Carninci *et al.*,[4] and the references therein.[12–15]

Materials. We routinely use the following materials to construct full-length cDNA libraries. Some alternatives are possible, especially relative to the cloning vectors and bacterial strains. Commonly used chemicals are not listed here. To avoid contamination due to "foreign" DNA and nucleases, all reagents for the preparation of cDNA libraries are stored separately from commonly used reagents.

Biological materials: Bacterial strain XL1 Blue *mrf'* (Stratagene, La Jolla, CA); cloning vector; λ ZapII (Stratagene).[10,11]

Enzymes and buffers: Restriction enzymes and reaction buffers, *Sst*I (GIBCO-BRL, Gaithersburg, MD) and *Xho*I [Takara (Ohtsu, Japan) or New England BioLabs (Beverly, MA)]; β-agarase and reac-

[7] P. S. Nelson, T. S. Papas, and C. W. Schweinfest, *Nucleic Acids Res.* **3,** 681 (1993).
[8] A. Kitamura and P. Carninci, unpublished observation (1997).
[9] J. Sambrook, E. F. Fritsch, and T. Maniatis, "Molecular Cloning," Cold Spring Harbor Laboratory Press, New York, 1989.
[10] J. M. Short, J. M. Fernandez, J. A. Sorge, and W. D. Huse, *Nucleic Acids Res.* **16,** 7583 (1992).
[11] J. M. Short and J. A. Sorge, *Methods Enzymol.* **216,** 495 (1992).
[12] K. Maruyama and S. Sugano, *Gene* **138,** 171 (1994).
[13] S. Kato, S. Sekine, S.-W. Oh, N.-S. Kim, Y. Umezawa, N. Abe, M. Yokoyama-Kobayashi, and T. Aoki, *Gene* **150,** 243 (1994).
[14] I. Ederly, L. L. Chu, N. Sonenberg, and J. Pelletier, *Mol. Cell. Biol.* **15,** 3363 (1995).
[15] CLONTECHniques technical bulletin, January (1996).

l with high efficiency even in the
l employs 5-methyl-dCTP instead
nthesis, the resulting cDNA is
iction digestion by the enzymes
e the unmethylated sites only on
:rs. The cDNAs is then cloned in
vectors are used here because of
-10 kb).[9] In addition, λ ZAPII is
pBluescript plasmid, which does
NA.[10,11]

using the powerful "thermostabi-
)are efficiently full-length cDNAs
lection of cDNA by biotinylated
:ngth cDNAs. Using these proto-
ibraries at high yield (several to
ı be routinely produced) without
which introduces sequence bias
clones in a library. For a more
cotols for full-length libraries, see
.n.[12-15]

wing materials to construct full-
s are possible, especially relative
ıins. Commonly used chemicals
:ion due to "foreign" DNA and
)n of cDNA libraries are stored

XL1 Blue *mrf'* (Stratagene, La
I (Stratagene).[10,11]
zymes and reaction buffers, *Sst*I
) and *Xho*I [Takara (Ohtsu, Ja-
verly, MA)]; β-agarase and reac-

Nucleic Acids Res. **3,** 681 (1993).
:ation (1997).
lolecular Cloning." Cold Spring Harbor

. Huse, *Nucleic Acids Res.* **16,** 7583 (1992).
:16, 495 (1992).
)4).
:awa, N. Abe, M. Yokoyama-Kobayashi.

er, *Mol. Cell. Biol.* **15,** 3363 (1995).
996).

tion buffer (New England BioLabs); T4 DNA ligation and reaction buffer (New England BioLabs); mouse mammary leukemia virus (MMLV) reverse transcriptase and reaction buffer (Superscript II; GIBCO-BRL); placental RNase inhibitor (Wako, Osaka, Japan); proteinase K (Sigma, St. Louis, MO) at 10 μg/μl in water, stored at −20° in aliquots; RNase I and reaction buffer (Promega, Madison, WI); terminal deoxynucleotidyltransferase (TdT; Takara); EX-*Taq* polymerase (Takara)

Chemicals and biochemicals: Low melting point agarose (Sea Plaque; FMC, Philadelphia, PA); RNase-free cetyltrimethylammonium bromide (CTAB; Sigma); deoxynucleotide triphosphates and 5-methyl-dCTP, sodium salts, stored at −80° as a 100 m*M* stock (Pharmacia, Piscataway, NJ); bovine serum albumin, DNase and RNase free (Takara); saturated trehalose [Fluka (Ronkonkoma, NY) or Sigma], approximately 80% (w/v), autoclaved; [α-^{32}P]dGTP, 3000 Ci/mmol (Amersham, Arlington Heights, IL); DE-81 chromatographic paper (Whatman, Clifton, NJ); $NaIO_4$ [Sigma or ICN (Irvine, CA)]; biotin hydrazide, long arm, also called biocytin hydrazide [Vector Laboratories (Burlingame, CA) or Sigma]; DNase- and RNase-free glycogen (Boehringer Mannheim, Indianapolis, IN); transfer RNA (tRNA, *E. coli*; Sigma). To ensure the absence of DNA, RNase-free DNase I (Promega) digestion is done, followed by SDS–proteinase K treatment, phenol–chloroform and chloroform extraction, and ethanol precipitation following standard procedures; first- and second-strand cDNA primers (see text for sequence and preparation)

Kits: λ DNA packaging extract (Max Plax; Epicentre, Madison, WI; mRNA extraction kit (Poly-A-Quick; Stratagene); magnetic porous glass (MPG) beads coated with streptavidin (CPG, Lincoln Park, NJ); CL-4B Spun column kit (Pharmacia)

Methods

λ *Cloning Vector Preparation*

Before starting the preparation of a cDNA library, prepare and test a high-efficiency, low-background cloning vector. As the cDNA fraction is radiolabeled, it becomes unstable and undergoes degradation if cloning is delayed. This will more dramatically affect the cloning of the longer than the shorter clones, resulting in a size bias and overrepresentation of shorter clones.

The oriented cDNA cloning protocol uses the λ ZAPII vector[10,11]; however, other vectors can be used provided that suitable restriction sites are present (see below). Here, alkaline phosphatase is not used, as it may dramatically decrease the cloning efficiency of the *Sst*I 5′ protruding ends. This drawback may be due to some exonuclease contamination, which is sometimes present in commercial sources of alkaline phosphatase. As we did not use alkaline phosphatase, there was no need for the formation of λ DNA concatamers by ligating the arms before the restriction, thus reducing the risk of mechanical shearing of the resulting very high molecular weight DNA.[9]

The highest quality restriction enzymes should be selected by checking the efficiency of religation after severalfold excess restriction digestion, as stated by the manufacturers.

Restrict the λ DNA in two steps, because the buffer requirements for *Sst*I and *Xho*I are different. First, cleave 15 μg of λ ZAPII in a large volume, such as 200 μl, by using 45 units of *Sst*I in the buffer provided by the manufacturer. Incubate for 2 hr at 37°. Notice that *Sst*I is an isoschizomer of *Sac*I, which cannot cleave hemimethylated cDNA. As it is difficult to separate the λ ZAPII arms by electrophoresis, a cut-check reaction should be done, by transferring 10 μl of the reaction to a separate tube containing 100 ng of a supercoiled plasmid in 0.5 μl of water. The plasmid should contain an *Sst*I restriction site (e.g., pBluescript; Stratagene). Incubate the cut-check reaction sample together with the main reaction. After incubating the sample for 2 hr, place the main reaction sample on ice and load the cut-check reaction sample on a 0.8% agarose minigel, together with a size marker and 100 ng of the starting uncut supercoiled plasmid. After electrophoresis, follow the extent of cleavage by checking the complete conversion from the supercoiled form of the plasmid to the linear form, which shows a slower migration.[9] If the reaction is incomplete, incubate the main reaction for an additional 1 or 2 hr and, if necessary, add more of the enzyme. When the reaction is complete, inactivate the *Sst*I at 65° for 10 min, change the buffer condition by adding 1.9 μl of 5 *M* NaCl and, after cooling the sample, add 45 units of *Xho*I. Repeat this procedure for the cut-check sample and incubate it for 2 hr at 37°. To reduce the background of the vector, a third cleavage can be performed at a third site on the λ DNA between *Xho*I and *Sst*I. For instance, we can add 30 units of *Not*I, the cut site of which lies between those of *Xho*I and *Sst*I in λ ZAPII, to the *Xho*I reaction. In this case, even if either *Xho*I or *Sst*I does not cleave completely, most of the termini that originate are incompatible for the ligation, thus reducing the background of the vector to a low level, usually less than 1–2%.

Finally, load the digested λ DNA in a 0.6% low melting point agarose minigel and run the electrophoresis for 60 min at 50 V. To follow the

ıses the λ ZAPII vector[10,11]; how-
that suitable restriction sites are
ısphatase is not used, as it may
cy of the *Sst*I 5′ protruding ends.
ıuclease contamination, which is
, of alkaline phosphatase. As we
as no need for the formation of
ɔefore the restriction, thus reduc-
ıe resulting very high molecular

:s should be selected by checking
ld excess restriction digestion, as

ause the buffer requirements for
e 15 μg of λ ZAPII in a large
of *Sst*I in the buffer provided by
Notice that *Sst*I is an isoschizomer
lated cDNA. As it is difficult to
resis, a cut-check reaction should
ion to a separate tube containing
ıl of water. The plasmid should
:script; Stratagene). Incubate the
e main reaction. After incubating
tion sample on ice and load the
ose minigel, together with a size
percoiled plasmid. After electro-
:hecking the complete conversion
to the linear form, which shows
ıplete, incubate the main reaction
, add more of the enzyme. When
stI at 65° for 10 min, change the
ıCl and, after cooling the sample,
ıre for the cut-check sample and
ɔackground of the vector, a third
e on the λ DNA between *Xho*I
its of *Not*I, the cut site of which
ZAPII, to the *Xho*I reaction. In
s not cleave completely, most of
ıle for the ligation, thus reducing
:el, usually less than 1–2%.
0.6% low melting point agarose
60 min at 50 V. To follow the

migration, instead of the usual high-energy 312-nm ultraviolet (UV) light transilluminator, use a mild 365-nm transilluminator, which is best if it is a hand-held type. Strong, short-wavelength UV light can cause DNA damage that will decrease the subsequent cloning efficiency, but a short exposure to 365-nm UV light will not. To minimize UV exposure, we usually do not take a picture of the gel at this stage and, instead, keep the previous cut-check data as a record of the experiment. The short stuffer is easily separated and the λ arms (which migrate together in the case of λ ZAPII) can be subsequently recovered from the gel by cutting with a sterile blade and transferral to a 2-ml Eppendorf tube. The tube is then briefly centrifuged to estimate the gel volume roughly, in order to calculate the number of units of β-agarase necessary to digest the agarose. To the gel, add the agarase buffer to the final 1× concentration and melt the gel at 65° for 10 min. Next, cool the gel to 40° for 5 min to equilibrate the temperature and add 3 units of β-agarase for each 100 μl of starting agarose gel. This concentration is in excess, but will completely remove all agarose, which would otherwise inhibit the ligation and the *in vitro* λ packaging reaction. Incubate the reaction for 2 to 4 hr at 40°. Finally, add 5 *M* NaCl to a final concentration of 1.2 *M*. The concentrated NaCl during the subsequent ethanol precipitation contributes to removal of the residual neutral polysaccharides.[16] Carefully perform phenol–chloroform and chloroform extraction (use a wide-bore pipette to avoid shearing of the DNA), then add 2 vol of ethanol. Incubate the reaction for 15 min on ice, then centrifuge at 12,000 rpm for 10 min to precipitate the DNA. Excessive centrifugation of λ DNA, which can lead to troublesome resuspension and the related risk of damage, should be avoided. After removing the supernatant, wash the pellet with 70% ethanol, centrifuge it for 2–3 min at 14,000 rpm, briefly air dry the pellet, and resuspend the DNA in 10–20 μl of 0.1× TE [1 m*M* Tris (pH 7.5 or 8.0) and 0.1 m*M* EDTA]. Measure the concentration by UV spectrophotometer and the polysaccharide contamination. A DNA preparation free of polysaccharides should have a 230/260 optical density (OD) ratio lower than 0.5, and a 260/280 OD ratio of approximately 1.8. Single-use aliquots (1–2 μg) of the vector can be conveniently stored for an indefinite period at −80°. Use one fraction of the vector to test the efficiency and background of the vector.

Testing Vector with Test Insert

Prepare a suitable test insert, for instance, any DNA of size 500–2000 bp cloned in pBluescript, by cleaving with *Sst*I and *Xho*I under standard

[16] G. Fang, S. Hammar, and R. Grumet, *Biotechniques* **13**, 52 (1992).

conditions. Purify the test insert from the agarose gel by standard silica-based methods,[17] or by using a kit such as GeneClean (Bio 101, La Jolla, CA) or Prep-A-Gene (Bio-Rad, Hercules, CA) as suggested by the manufacturers, or by adapting the preceding protocol using β-agarase.

Set up the following ligation reactions in a final volume of 5 μl:

λ DNA and test insert:

Reaction a: Only purified λ vector (100 ng)

Reaction b: λ vector (100 ng) and test insert, ratio 1:3

Reaction c: λ vector (100 ng) and test insert, ratio 1:1

Reaction d: λ vector (100 ng) and test insert, ratio 3:1

DNA ligase buffer: 0.5 μl

DNA ligase (200 Weiss units/μl): 0.4 μl

Incubate the reactions overnight at 16°. On the same day, inoculate a single colony of XL1-Blue *mrf'* in LB broth–0.2% maltose–10 m*M* $MgSO_4$, and incubate at 37° for 6–8 hr to measure the efficiency on the next day. Prepare the bacterial cells using standard procedures.[9] The next day, package half (2.5 μl) of the ligation reactions. At the same time, package 50 ng of the cut, but not religated, λ DNA (tube e), and 50 ng of the starting noncleaved λ DNA (tube f) in a 2.5-μl volume.

Thaw one Max Plax packaging extract[18] quickly by holding it between the fingers and promptly add 7.5 μl of the extract to each aliquot (tubes a–f). Conduct the reaction at 22° for 1.5 to 2 hr, then add 500 μl of SM buffer[9] and 20 μl of chloroform ($CHCl_3$) and titer as described by Sambrook *et al.*[9] After 6 to 8 hr, count the plaques and calculate the relative titer expected for 1 μg of DNA.

A satisfactory vector may give about 10^7 PFU/μg of λ DNA for the positive reactions b and c. The cloning background is given by reaction a. The titer of this reaction should be less than 5% for an acceptably low background when using the cDNA. As a reference value, a typical titer for the positive control is around 10^8 PFU/μg of λ DNA (reaction f), and less than 10^4 PFU/μg of λ DNA for the cleaved unligated sample (reaction e).

Working with RNA

For details on setting up an RNase-free environment, refer to more detailed laboratory guides.[9,19] As a rule, when working with RNA samples, gloves should be worn at all times; it is best to work off a bench dedicated to RNA work. Plasticware and solutions should be autoclaved (except

[17] R. Boom, C. J. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-vanDillen, and J. van der Noordaa, *J. Clin. Micribiol.* **28**, 495 (1990).

[18] E. J. Gunther, N. E. Murray, and P. Glazer, *Nucleic Acids Res.* **21**, 3903 (1993).

[19] D. D. Blumberg, *Methods Enzymol.* **152**, 20 (1987).

e agarose gel by standard silica-
ıs GeneClean (Bio 101, La Jolla,
:, CA) as suggested by the manu-
rotocol using β-agarase.
: in a final volume of 5 μl:

· (100 ng)
test insert, ratio 1:3
test insert, ratio 1:1
test insert, ratio 3:1

l μl
6°. On the same day, inoculate a
:h–0.2% maltose–10 m*M* $MgSO_4$,
·e the efficiency on the next day.
procedures.[9] The next day, pack-
:. At the same time, package 50
:ube e), and 50 ng of the starting
olume.
t[18] quickly by holding it between
he extract to each aliquot (tubes
to 2 hr, then add 500 μl of SM
ıd titer as described by Sambrook
: and calculate the relative titer

10^7 PFU/μg of λ DNA for the
ıckground is given by reaction a.
than 5% for an acceptably low
·eference value, a typical titer for
: of λ DNA (reaction f), and less
:d unligated sample (reaction e).

ree environment, refer to more
hen working with RNA samples,
:st to work off a bench dedicated
·s should be autoclaved (except

ınsen, P. M. E. Wertheim-vanDillen, and
)90).
leic Acids Res. **21,** 3903 (1993).
7).

for solution D, phenol–chloroform, CTAB–urea, and 7 *M* guanidinium chloride solutions; see below), and glassware should be baked at 250° before use. Even if not specified, all reagents should be RNase free, at least until the end of the biotinylation of the mRNA–cDNA hybrid (see below). From our experience, using diethylpyrocarbonate (DEPC) is not recommended because autoclaving is sufficient to remove any RNase contamination if dedicated materials are used. We did find that the RNA was sometimes degraded after incubation at 65° in DEPC-treated water, which has a moderately acidic pH. To prepare RNase-free water, autoclaved double-distilled water or water of higher quality (for instance, MilliQ; Millipore, Bedford, MA), without DEPC treatment, is usually acceptable.

RNA Preparation

Total RNA is commonly prepared from mouse tissues by modification of a standard procedure[20] with adaption of the CTAB precipitation method[21,22] for selective removal of polysaccharides. This protocol also has been used for cultured cells and tissues of other mammals. We have found that the presence of polysaccharides, which may be subsequently biotinylated, inhibits the binding of full-length cDNA to streptavidin-coated magnetic beads (see below), perhaps by competing for the available streptavidin sites. Under the conditions described, the cationic micelles formed by CTAB undergo selective formation of complexes by charging the anionic nucleic acids, thus leaving in solution all neutral polysaccharides and residual proteins, which are easily eliminated in a single step.

This protocol is suitable for approximately 0.5–1 g of tissue. Volumes should be scaled up or down if the amount of tissue differs considerably. Quickly homogenize the fresh tissue in 10 ml of solution D [4 *M* guanidinium thiocyanate, 25 m*M* sodium citrate (pH 7.0), 100 m*M* 2-mercaptoethanol, 0.5% *N*-laurylsarcosine],[20] followed by the addition of 1 ml of 2 *M* sodium acetate, pH 4.0, 8 ml of water-equilibrated phenol,[9,20] and 2 ml of chloroform. High-quality RNA is obtained at consistently high yield with freshly equilibrated phenol (no earlier than a few weeks before use) or from frozen aliquots. After 15 min on ice, samples are centrifuged at 7500*g* for 15 min. The upper, aqueous phase containing the RNA is then gently separated and transferred into a new tube by pipetting, taking care to avoid the precipitated material at the layer between the two phases, which contains genomic DNA and proteins. Next, precipitate the RNA from the aqueous

[20] P. Chomczynski and N. Sacchi, *Anal. Biochem.* **162,** 156 (1987).
[21] G. Del Sal, G. Manfioletti, and C. Schneider, *BioTechniques* **7,** 514 (1989).
[22] S. Gustincich, P. Carninci, G. Del Sal, G. Manfioletti, and C. Schneider, *Biotechniques* **11,** 298 (1991).

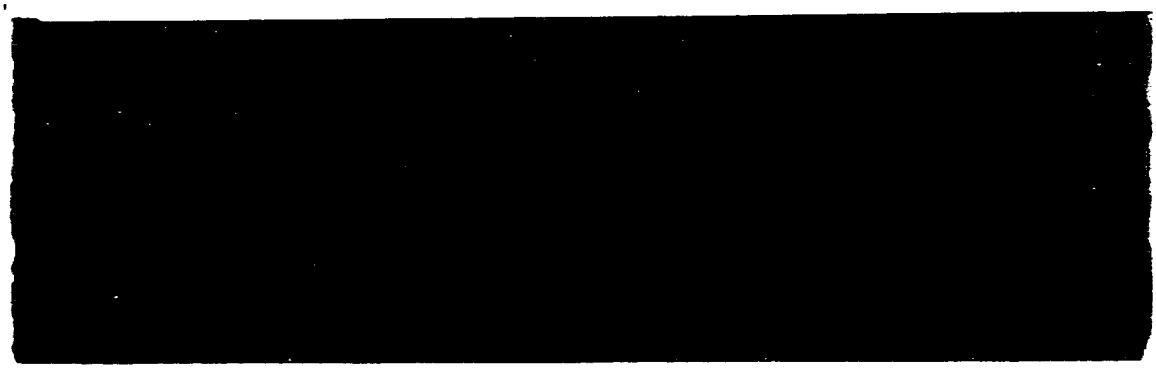
phase by adding 1 vol of isopropanol. The sample is then incubated for 1 hr on ice, after which the RNA is pelleted by centrifugation at 7500*g* for 15 min. The pellet is washed twice with 70% ethanol, each time followed by centrifugation at 7500 rpm for 2 min, in order to remove the SCN salts, which would interfere with the following CTAB precipitation by forming a chemical precipitate. Selective CTAB precipitation of mRNA is performed after complete RNA resuspension in 4 ml of water. Subsequently, 1.3 ml of 5 *M* NaCl is added and the RNA is selectively precipitated by adding 16 ml of a CTAB–urea solution [1% CTAB, 4 *M* Urea, 50 m*M* Tris (pH 7.0), 1 m*M* EDTA (pH 8.0)]. The CTAB solution should not be autoclaved but prepared with RNase-free reagents (RNase-free CTAB is available from Sigma), water, and glassware. After 15 min of centrifugation at 7500 rpm at room temperature, the RNA is resuspended in 4 ml of 7 *M* guanidine chloride. The resuspension of RNA using a high salt concentration is necessary to remove the CTAB by ionic exchange.[21,22] Resuspended RNA is finally precipitated by adding 8 ml of ethanol, and after 1 hr on ice, the RNA is pelleted by centrifuging at 7500 rpm for 15 min. The pellet is washed with 70% ethanol and resuspended in water. RNA purity is monitored by reading the OD ratio at 230, 260, and 280 nm. Removal of polysaccharides is successful when the 230/260 ratio is lower than 0.5, and effective removal of proteins is obtained when the 260/280 ratio is higher than 1.8 and possibly ~2.0.

The messenger RNA is subsequently prepared by using commercial kits based on oligo(dT)-cellulose and starting from the guanidinium isothiocyanate–CTAB-purified RNA.[9] We have found Poly-A-Quick (Stratagene) to provide a satisfactory yield of mRNA under the conditions recommended by the manufacturer. For the latter use, we redissolve the poly(A)$^+$ RNA at a high concentration of 1 to 2 μg/μl (see [3] in this volume).[22a] From 500 μg of total RNA, 10 to 20 μg of mRNA can be routinely purified. Only one oligo(dT) selection of the mRNA is done; extensive pretreatment of the samples may cause underrepresentation of long mRNAs. Moreover, traces of cDNA derived from ribosomal RNAs (that lack the cap structure) are lost during selection of the full-length cDNA.

Preparation of Primers

First-strand cDNA primers should be thoroughly purified and rendered RNase free. Primer adapters containing the *Xho*I site [5′$(GA)_8$ ACTCGAG$(T)_{16}$VN-3′] are synthesized by adding V as a degenerate base in synthesis as G, A, or C, and N as G, A, T, or C, by synthesizing four

[22a] M. Liu, Y. V. B. K. Subramanyam, and N. Baskaran, *Methods Enzymol.* **303**, [3], 1999 (this volume).



he sample is then incubated for
ed by centrifugation at 7500*g* for
70% ethanol, each time followed
n order to remove the SCN salts.
TAB precipitation by forming a
ipitation of mRNA is performed
l of water. Subsequently, 1.3 ml
lectively precipitated by adding
AB, 4 *M* Urea, 50 m*M* Tris (pH
olution should not be autoclaved
(RNase-free CTAB is available
15 min of centrifugation at 7500
ıspended in 4 ml of 7 *M* guanidine
a high salt concentration is neces-
ıange.[21,22] Resuspended RNA is
ıanol, and after 1 hr on ice, the
n for 15 min. The pellet is washed
er. RNA purity is monitored by
nm. Removal of polysaccharides
r than 0.5, and effective removal
ratio is higher than 1.8 and possi-

prepared by using commercial
ng from the guanidinium isothio-
ound Poly-A-Quick (Stratagene)
der the conditions recommended
ve redissolve the poly(A)$^+$ RNA
see [3] in this volume).[22a] From
A can be routinely purified. Only
done; extensive pretreatment of
ion of long mRNAs. Moreover,
NAs (that lack the cap structure)
ı cDNA.

e thoroughly purified and ren-
ıtaining the *Xho*I site [5′(GA)$_8$
y adding V as a degenerate base
A, T, or C, by synthesizing four

karan. *Methods Enzymol.* **303**, [3], 1999

different oligonucleotides from the four different CPG matrices. After ammonia deblocking, oligonucleotides are precipitated twice with 10 vol of butanol,[23] and further purified by acrylamide gel electrophoresis following standard techniques.[9] We use 12% acrylamide, 8 *M* urea, and 1× TBE (Tris–borate–EDTA). Alternatively, primers can be purified by high-performance liquid chromatography (HPLC). To ensure that purified primers are RNase free, the primers should be extracted with phenol–chloroform and chloroform and precipitated by addition of 0.2 *M* NaCl and 2.5 vol of ethanol, incubated at −20° for 1 hr, and centrifuged for 20 min at 15,000 rpm. After washing the pellet with 80% ethanol, primers are resuspended in water at a high concentration (>2 μg/μl). After checking the OD and mixing together the four primers in equal parts, the primer mixture is ready to prime the first-strand cDNA synthesis.

cDNA Preparation: First-Strand Synthesis

The average size of the first-strand cDNA when the RT is thermostabilized by addition of trehalose is greater than that of cDNA synthesized under standard conditions. The highest performance of the thermostabilized RT is obtained with temperature cycling between 55 and 60°.[2] Before starting the reaction, a thermal cycler with a hot lid (e.g., MJ Research, Watertown, MA) is set with the following first-strand cDNA synthesis program:

Step 1: 45° for 2 min (hot start)
Step 2: Negative ramp: go to 35° in 1 min (gradient annealing)
Step 3: 35° for 2 min (complete annealing)
Step 4: 45° for 5 min
Step 5: Positive ramp: +15° (until 60°) at +0.1°/sec
Step 6: 55° for 2 min
Step 7: 60° for 2 min
Step 8: Go to step 6 for 10 additional times
Step 9: +4° forever

The total time required is about 60 min. The ramp from 45 to 35° is designed specifically to anneal the primer at the beginning of the poly(A) of the mRNA.

To prepare the first-strand cDNA, put together the following reagents in three different 0.5-ml PCR tubes (A, B, and C):

Tube A: In a final volume of 24 μl, add the following:

mRNA:	5 to 10 μg
First-strand primer mixture:	5 μg
Glycerol (80%):	11.2 μl

[23] M. Sawadogo and M. W. Van Dyke, *Nucleic Acids Res.* **19**, 674 (1991).

The concentration of primers should be scaled up or down depending on the size, if different primer adapters are used. See below for a discussion of the possible primer adapter sites that can be used in the oligo primers. Heat the mixture (mRNA, primer, and glycerol) at 65° for 10 min to dissolve the secondary structures of mRNA. During the incubation, quickly prepare the reagent mix (tube B) and tube C:

Tube B: In a final volume of 76 μl, add the following:

First-strand buffer (5×):	18.2 μl
Dithiothreitol (DTT, 0.1 *M*):	9.1 μl
dATP, dTTP, dGTP, and 5-methyl-dCTP (instead of dCTP), 10 m*M* each:	6.0 μl
Bovine serum albumin (BSA, 2.5 γ/λ):	2.3 μl
Saturated trehalose (approximately 80%):	29.6 μl
Placental RNase inhibitor:	1.0 μl
Superscript II reverse transcriptase (200 U/μl):	10.0 μl

In tube C, place 1.0 μl of [α-^{32}P]dGTP or, alternatively, [α-^{32}P]dTTP or [α-^{32}P]dATP. Do not use [α-^{32}P]dCTP as a tracer; hemimethylated cDNA containing a fraction of unmethylated C becomes partially sensitive to the restriction digestion, thus leading to internal cleavage of the cDNA before cloning. If the mixture in tube B is not ready, after incubation at 65°, tube A can be left on ice for only a short time, to minimize the reformation of secondary structures of RNA.

Quickly put tubes A, B, and C on the thermal cycler to begin step 1, the hot-start incubation at 45°. After 15 sec, to equilibrate the temperature, transfer the contents of tube B to A, mix quickly but thoroughly, and transfer 25 μl of the resulting A + B mixture into tube C and mix. Complete the manipulation before the beginning of step 2 for an exact estimation of the incorporation of [α-^{32}P]dGTP that reflects the yield of the synthesized cDNA. Alternatively, the contents of tubes A, B, and C can be quickly mixed on ice instead of on the thermal cycler, and transferred immediately to 45° to begin step 1 of the cycling program. We usually label only 25% of the cDNA; the remaining 75% is unlabeled and thus does not undergo radiodegradation even if cloning is delayed.

At the end of the reaction (step 9), take 0.5 μl of the reaction in tube C and spot it onto a small square of DE-81 paper; keep a 0.5-μl aliquot in a separate tube for subsequent alkaline gel analysis. Measure the radioactivity of the spot before and after three 10-min washings with 50 ml of 0.5 *M* sodium phosphate, pH 7.0, followed by a brief washing with water, and 70% ethanol and quick air drying. Typical incorporation rates range between 3

ɔe scaled up or down depending
: used. See below for a discussion
:an be used in the oligo primers.
:erol) at 65° for 10 min to dissolve
g the incubation, quickly prepare

ld the following:

	18.2 μl
	9.1 μl
CTP	
	6.0 μl
):	2.3 μl
0%):	29.6 μl
	1.0 μl
00 U/μl):	10.0 μl

'P or, alternatively, [α-^{32}P]dTTP
ΓP as a tracer; hemimethylated
'ted C becomes partially sensitive
o internal cleavage of the cDNA
is not ready, after incubation at
a short time, to minimize the
NA.
ɔ thermal cycler to begin step 1,
:, to equilibrate the temperature,
ıix quickly but thoroughly, and
re into tube C and mix. Complete
step 2 for an exact estimation of
lects the yield of the synthesized
ɔes A, B, and C can be quickly
:ler, and transferred immediately
ram. We usually label only 25%
ɔeled and thus does not undergo
ɔd.
:ke 0.5 μl of the reaction in tube
31 paper; keep a 0.5-μl aliquot in
analysis. Measure the radioactiv-
ıin washings with 50 ml of 0.5 *M*
'rief washing with water, and 70%
ɔrporation rates range between 3

to 8%, depending on the actual concentration of mRNA, which might have been overestimated by OD reading owing to contamination of ribosomal RNA. Alternatively, some trouble may have occurred in the first-strand synthesis, mainly owing to RNase contamination or decreased enzymatic activity due to the mishandling of the RT. In the first case, the cDNA, if analyzed by alkaline gel, will have an average size of 2 kb or higher, and the longest cDNA will be longer than 10 kb. In the second case, the cDNA will be notably shorter than 2 kb. In this case, the procedure should be stopped and the reagents tested.

Calculate the total yield (μg) of the synthesized cDNA, starting from the quantity of the dNTPs used in the reaction. The four dNTPs in the stock solution are present at 10 m*M* each, for a total concentration of 40 m*M*. The concentration of the tracer, [α-^{32}P]dGTP, is not relevant to this calculation. Because 6 μl is used in the reaction, it gives 6×10^{-6} (liter) × 40×10^{-3} (moles/liter) = 2.4×10^{-7} mol of dNTPs in the reaction. Because the average molecular weight of a residue is 340, in the reaction there are $340 \times 2.4 \times 10^{-7} = 8.17 \times 10^{-5}$ g of dNTPs, that is 81.7 μg. Consequently, the final yield is given by

81.7 μg × (% of incorporation)/100 = micrograms of first-strand cDNA synthesized

Organic Phase Extraction and cDNA Precipitation. Transfer both the "hot" and "cold" first-strand synthesis (tubes B and C) to a centrifuge tube (PCR tubes cannot be used for the centrifugation at 15,000 rpm with organic solvents, because they may be crushed) and add

EDTA (0.5 *M*) to a final concentration of 10 m*M* (2 μl)
Sodium dodecyl sulfate (SDS, 10%) to a final concentration of 0.2% (2 μl)
Proteinase K (10 μg/μl) to a final concentration of 100 ng/μl (1 μl)

Incubate at 45° for at least 15 min. Although the proteinase K incubation is time consuming, it contributes greatly to increased sample purity. The proteins (especially BSA) must be completely removed to facilitate subsequent redissolution of the pellet. If this is not done, the cDNA becomes "sticky." Moreover, in the absence of proteinase K treatment, cDNA is partially entrapped in the precipitated material that is sometimes observed in the layer between the organic and aqueous phase (see below), which appears "dirty" and contains some cDNA (radioactive) that cannot be back extracted.

Finally, add, in a final volume of 200 μl:

Ammonium acetate (10 *M*):	60 μl
Water	37 μl

Perform phenol–chloroform and chloroform extraction and back extraction:

1. Add 0.5 vol (100 μl) of phenol.
2. Add 0.5 vol (100 μl) of chloroform.
3. Vortex moderately until the two phases mix.
4. Leave on ice for 1–2 min.
5. Centrifuge for 2 min at 15,000 rpm.
6. Carefully remove the aqueous phase (upper layer), using a P-200 pipette, and transfer to a fresh 1.5-ml tube. Transfer as much as possible of the upper phase but avoid touching or transferring any contaminant material from the lower, organic phase. Keep the tube with phenol for the back extraction (see below).
7. Mix the phenol–chloroform-extracted cDNA and 200 μl of chloroform.
8. Vortex gently.
9. Centrifuge for 2 min at 15,000 rpm.
10. Transfer the aqueous phase (with a P-200 pipette) to a fresh, clean tube. Keep the tube with chloroform for the back extraction.

Perform back extraction to recover the residual cDNA from the residual aqueous phase that could not be recovered after the first extraction. This procedure helps increase the yield by 10–15% at each step.

1. Add 50 μl of water to the phenol–chloroform tube.
2. Vortex gently.
3. Centrifuge again for 2 min at 15,000 rpm.
4. Transfer the upper phase to the chloroform tube; discard the phenol–chloroform organic phase.
5. Vortex and centrifuge for 2 min at 15,000 rpm.
6. Transfer the upper phase to the chloroform tube.
7. Vortex and centrifuge for 2 min at 15,000 rpm.
8. Transfer the upper, aqueous phase to the previously extracted fraction of cDNA.

Checking the efficiency of cDNA recovery by hand-held monitor is not particularly effective at this stage owing to the large excess of free nucleotides. This should be done instead for all subsequent extraction/back extraction procedures, in the absence of free radioactive nucleotides. Refer to this protocol (except for the addition of ammonium acetate) for proteinase K incubation and extraction/back extraction at subsequent stages.

Finally, precipitate the cDNA with 2.5 vol of absolute ethanol.

1. Add 625 μl of absolute ethanol.
2. Incubate at $-80°$ for 20 min, or at $-20°$ or on ice for 30–60 min.

The ethanol precipitation step is always a convenient stopping point; the procedure can be interrupted for hours or overnight, if necessary.

After the incubation, obtain the cDNA by centrifugation at 15,000 rpm for 15 min. Remove the supernatant (*caution:* very hot) and wash the pellet twice with 800 μl of 80% ethanol. Each time, add 80% ethanol to the tube on the opposite side from where the cDNA is being pelleted and centrifuge for 2–3 min at 15,000 rpm. Finally, resuspend the cDNA in 47 μl of water. The extent of resuspension can be determined by checking the counts in the tube and in the solution by hand-held monitor. When the resuspension is complete, less than 5% of the total count will remain on the tube wall. Check that there are no clusters of nonredissolved cDNA in the solution. Do not use TE (Tris–EDTA) buffer at this stage to resuspend the cDNA, because the polyhydroxyl groups of TE are oxidized in the subsequent step and may quench the biotinylation reaction.

Biotinylation Reaction

OXIDATION OF DIOL GROUPS OF mRNA. In a final volume of 50 μl, add the following:

Resuspended cDNA sample
Sodium acetate buffer, pH 4.5 (1 *M*): 3.3 μl
Freshly prepared solution of $NaIO_4$ to a final concentration of 5 m*M*

Incubate on ice in the dark for 45 min. A note of caution: as exposure to light and high temperature makes the oxidation steps less specific, wrap the tube with aluminum foil. Finally, precipitate the cDNA by adding

SDS (10%): 0.5 μl
NaCl: 11 μl
Isopropanol: 61 μl

Incubate on ice for 45 min, or at -20 to $-80°$ for 30 min, in the dark. Centrifuge for 10 min at 15,000 rpm. Rinse the pellet twice with 70% ethanol, and each time recentrifuge at 15,000 rpm for 2 min. Resuspend the cDNA-pellet in 50 μl of water. The resuspension can be partially inhibited by the acidic pH of the acetate buffer used before the precipitation.

DERIVATIZATION OF OXIDIZED DIOL GROUPS. Prepare 10 m*M* biotin hydrazide long arm (MW 371.51) in water. Note that this reagent requires a long time and extensive mixing for complete solubilization in water. Always use fresh solutions of biotin hydrazide; frozen aliquots are sometimes not reactive.

To the cDNA add, in a final volume of 210 μl:

Sodium acetate buffer, pH 6.1 (1 *M*):	5 μl
SDS (10%):	5 μl
Biotin hydrazide long arm (10 m*M*):	150 μl

Incubate overnight (10–16 hr) at room temperature (22–26°). On the next day, to precipitate the biotinylated cDNA, add

Sodium acetate, pH 6.1 (1 *M*):	75 μl
NaCl (5 *M*):	5 μl
Absolute ethanol:	750 μl

Incubate on ice for 1 hr or at −80 to −20° for 30 min. Centrifuge the sample at 15,000 rpm for 10 min. Wash the precipitate once with 70% ethanol and once with 80% ethanol. Accurate washing of the pellet will help to remove the free biotin hydrazide. If not removed, it will compete with the biotinylated cDNA to bind the streptavidin beads. Redissolve the sample in 70 μl of 0.1× TE [1 m*M* Tris (pH 7.5), 0.1 m*M* EDTA]. Monitor for complete resuspension using a hand-held monitor.

Magnetic Bead Preparation. Streptavidin-coated porous glass beads (MPG) were selected because of the low nonspecific binding of the nucleic acids to the glass matrix compared with latex magnetic beads, and because of their excellent binding capacity.[24] To minimize further the nonspecific binding of nucleic acids, the beads are further preincubated with DNA-free tRNA. Start the blocking of the magnetic beads with DNA-free tRNA, and subsequent washings, just prior to the RNase I treatment as described below. Careful washing at this stage will remove all free streptavidin, which may have been released from the beads. Any trace of free streptavidin will decrease the efficiency of cDNA capture. Once washed, the beads should be used within 1 hr. Streptavidin does not remain stable longer than this in the BSA-containing storage buffer.

1. To 500 μl of streptavidin-coated MPG beads add 100 μg of DNA-free tRNA (at a concentration of 10–50 μg/μl).
2. Incubate on ice for 30 min with occasional mixing.
3. Separate the beads with a magnetic stand (for 3 min) and remove the supernatant.
4. Wash three times with 500 μl of washing/binding solution [2 *M* NaCl, 50 m*M* EDTA (pH 8.0)].

Washings, here and later, are performed by gentle redissolving of the beads after their capture by the appropriate wash solution, followed by their recapture for 3 min on a magnetic stand and removal of the solution

[24] P. Carninci, unpublished observation (1995).

of 210 μl:
5 μl
5 μl
150 μl
m temperature (22–26°). On the
cDNA, add
l
l
l
−20° for 30 min. Centrifuge the
the precipitate once with 70%
curate washing of the pellet will
. If not removed, it will compete
treptavidin beads. Redissolve the
H 7.5), 0.1 m*M* EDTA]. Monitor
ield monitor.
idin-coated porous glass beads
nonspecific binding of the nucleic
ıtex magnetic beads, and because
minimize further the nonspecific
urther preincubated with DNA-
ıetic beads with DNA-free tRNA,
RNase I treatment as described
move all free streptavidin, which
Any trace of free streptavidin will
. Once washed, the beads should
ot remain stable longer than this

IPG beads add 100 μg of DNA-
10–50 μg/μl).
casional mixing.
ic stand (for 3 min) and remove

shing/binding solution [2 *M* NaCl,

ned by gentle redissolving of the
riate wash solution, followed by
stand and removal of the solution

by pipetting. Vortexing should be gentle, especially during the capture and washing of the cDNA.

RNase I Treatment. RNase I was selected because it is the only enzyme in this category that cleaves RNA at any base in a sequence-independent manner.[25] Consequently, no RNase cocktail is needed for the cleavage of any sequence. Also, RNase I can be inactivated simply by discarding it in solutions containing SDS at a concentration of at least 0.1% or higher. Thus, RNase I can be quite safely used in a RNA-dedicated laboratory.

To the cDNA sample (70 μl) add, in a final volume of 200 μl:
RNase I buffer (Promega): 20 μl
RNase I (5 U/μl; Promega): 200 units; glycerol concentration up to 10% does not cause a problem

Incubate at 37° for 15 min. Avoid prolonged incubation that may decrease the yield. To stop the reaction, put the sample on ice and add 100 μg of tRNA and 100μl of 5 *M* NaCl. The addition of tRNA at this stage prevents the nonspecific binding of cDNA to the beads (see the next section).

Capture of Full-Length cDNA. cDNA capture is performed in a high sodium chloride concentration. High-salt buffers help the binding of long cDNAs[26] and minimize the nonspecific interaction of nucleic acids with streptavidin. To capture the full-length cDNA, mix the RNase I-treated cDNA and the washed beads as follows:

1. Resuspend the beads in 400 μl of the washing/binding solution.
2. Transfer the beads into the tube containing the biotinylated first-strand cDNA.
3. After mixing, gently rotate the tube for 30 min at room temperature.

Full-length cDNA remains on the beads, but the shortened cDNAs do not. Separate the beads from the supernatant on a magnetic stirrer. The uncaptured fraction can be precipitated with 0.6 vol of isopropanol for further analysis by alkaline gel (optional step).

Washing Beads. Gently wash the beads in the indicated buffer to remove the nonspecifically absorbed cDNAs.

1. Twice with washing/binding solution
2. Once with 0.4% SDS, 50 μg/ml tRNA
3. Once with 10 m*M* Tris-HCl (pH 7.5), 0.2 m*M* EDTA, 40 μg/ml tRNA, 10 m*M* NaCl, 20% glycerol
4. Once with 50 μg/ml tRNA in water

[25] J. Meador III, B. Cannon, V. J. Cannistraro, and D. Kennel, *Eur. J. Biochem.* **187,** 549 (1990).
[26] S. T. Kostopulos and A. P. Shuber, *Biotechniques* **20,** 199 (1996).

cDNA Release from Beads. After the washings, release the cDNA from the beads by alkali treatment, which denatures the cDNA–mRNA hybrid and hydrolyzes the mRNA. To the tube containing the beads, add 50 μl of 50 m*M* NaOH and 5 m*M* EDTA, then briefly stir and incubate 10 min at room temperature with occasional mixing. Separate the magnetic beads and transfer the eluted cDNA on ice into a separate tube containing 50 μl of 1 *M* Tris-HCl, pH 7.5. Repeat the elution cycle with 50 μl of 50 m*M* NaOH, 5 m*M* EDTA once or twice until most of the cDNA, 80–90% as measured by monitoring the counts per minute with a hand-held monitor, can be recovered from the beads. Pool the eluted fractions.

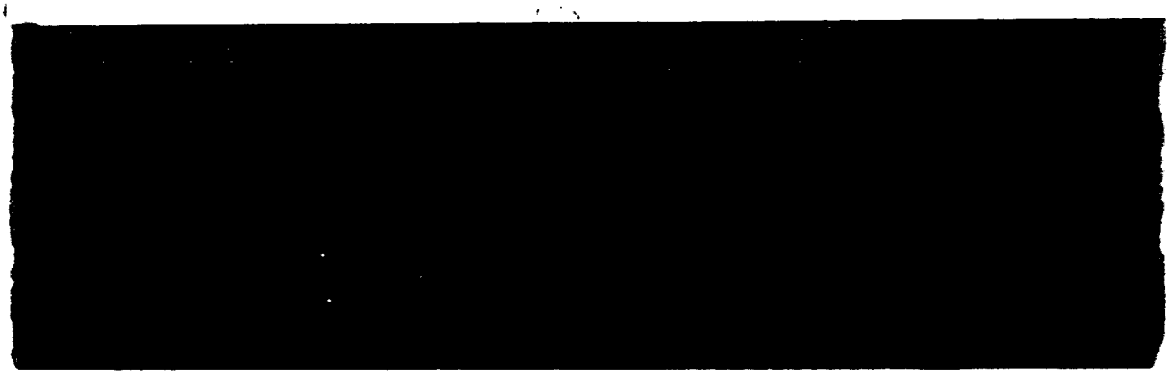
As done for the first-strand reaction, treat the cDNA with proteinase K and phenol–chloroform extraction, including back extraction. At this stage, chloroform extraction is not necessary. All the precipitations are carried out from this step onward by using 0.2 *M* sodium chloride instead of 2 *M* ammonium acetate. To increase the efficiency of the precipitation, add 3 μg of glycogen and precipitate the sample with 1 vol of isopropanol, following standard procedures, then wash with 70% ethanol. Resuspend the pellet in 50 μl of 0.1× TE. Check the resuspension of cDNA by hand-held monitor and visual inspection for the presence of clusters of nonresuspended cDNA.

Siliconized tubes may be used to minimize the adsorption of the ssDNA to the plastic wall. The presence of carrier and the high purity of cDNA after proteinase K digestion and organic extraction can solve most of the trouble encountered due to the stickiness of single-stranded cDNA, which can usually be redissolved completely. To prevent the cDNA from sticking to the tube, do not dry the pellet under vacuum.

CL-4B Spun Column Fractionation of cDNA. Prior to the tailing reaction, traces of the first-strand primer, which become nonspecifically absorbed to the beads and later released, should be removed by gel filtration. The protocol described here uses a Sepharose CL-4B spun column (Pharmacia), essentially as described by the manufacturer. Depending on the experience and choice of the operator, spinning the column may be substituted with gravity column separation. We also successfully tested matrices other than CL-4B, such as S400.

For the spun column protocol:

1. Shake the column several times (until the matrix is completely redissolved), then let it stand upright.
2. Remove the upper cap, then remove the bottom one.
3. Drain the buffer from the column. If air bubbles enter the column, add column storage buffer again and go back to step 1.



washings, release the cDNA from
atures the cDNA–mRNA hybrid
containing the beads, add 50 μl
briefly stir and incubate 10 min
ing. Separate the magnetic beads
a separate tube containing 50 μl
ition cycle with 50 μl of 50 m*M*
l most of the cDNA, 80–90% as
ninute with a hand-held monitor,
ie eluted fractions.
treat the cDNA with proteinase
cluding back extraction. At this
ssary. All the precipitations are
g 0.2 *M* sodium chloride instead
he efficiency of the precipitation,
sample with 1 vol of isopropanol,
h with 70% ethanol. Resuspend
resuspension of cDNA by hand-
presence of clusters of nonresus-

nize the adsorption of the ssDNA
er and the high purity of cDNA
extraction can solve most of the
of single-stranded cDNA, which
prevent the cDNA from sticking
vacuum.
cDNA. Prior to the tailing reac-
hich become nonspecifically ab-
ould be removed by gel filtration.
rose CL-4B spun column (Phar-
nanufacturer. Depending on the
inning the column may be substi-
also successfully tested matrices

til the matrix is completely redis-

ve the bottom one.
If air bubbles enter the column,
nd go back to step 1.

4. Apply 2 ml of the buffer [10 m*M* Tris (pH 8.0), 1 m*M* EDTA, 0.1% SDS, and 100 m*M* NaCl] and drain twice by gravity.
5. Put the column into a 15-ml centrifuge tube, then centrifuge at 400*g* for 2 min in a swing-out rotor at room temperature. If cracks appear in the matrix, suspend it in the buffer defined in step 4; shake and repeat the procedure from step 4.
6. Apply 100 μl of buffer to the column, then centrifuge at 400*g* for 2 min. Check the eluted volume. If it is different from the input volume (100 μl), repeat this step until the eluted volume is the same as the added volume.
7. Set a 1.5-ml tube (after cutting off the cap) into the 15-ml centrifuge tube, then apply the sample to the column. Centrifuge at 400*g* for 2 min.
8. Collect the eluted fraction in a separate tube. Apply 50 μl of buffer to the column, repeat the centrifugation, and collect the fraction in a separate tube.
9. Repeat step 8 three to five more times; keep the eluted fractions separate.

Collected fractions should be counted by scintillation. The bulk of the cDNA (70% or more of the counts) is usually in the first three or four fractions, which contain the longer cDNAs and are free of contaminating first-strand primer. The subsequent fractions contain only short cDNAs and first-strand primer and should be discarded to prevent introducing short cDNAs and first strand–second primer dimers into the library. Pool the selected fractions (usually three or four), add 1 μg of carrier tRNA and NaCl to a final concentration of 0.2 *M*, and precipitate the cDNA by adding 1 vol of isopropanol. After precipitation and washing, dissolve the pellet in 31 μl of water.

Oligo(dG) Tailing of First-Strand cDNA. Oligo(dG) tailing is adopted here owing to its proven capacity to clone the 5′ ends of cDNAs from the first transcribed base.[27] Deoxyguanosine triphosphate is the selected nucleotide, because the reaction stops after the addition of approximately 15 G residues. Other methods to provide a sequence suitable for priming the second-strand synthesis, such as the RNA ligase-mediated protocols, yield too low a level of efficiency and sequence bias to be used in this strategy.[12,13]

To check the tailing reaction, prepare in a separate tube 0.5 ng of an oligonucleotide 20–40 bases long in a volume of 0.5 μl. This oligonucleotide

[27] P. L. Deininger, *Methods Enzymol.* **152**, 371 (1987).

should be 5′ end labeled with ^{32}P by polynucleotide kinase (PNK), following standard procedures, extracted with phenol–chloroform and chloroform, and precipitated with ethanol by using 2 μg of glycogen as a carrier.[9] Dilute the oligonucleotide at a concentration of approximately 1 ng/μl. The labeled oligonucleotide should be stored at −20° and can be used for no more than 4 weeks after labeling. The function of this oligonucleotide, the sequence of which is not really important unless there are no strong internal secondary structures, is to check the efficiency and length of the G tail in a parallel reaction.

Just prior to tailing, the cDNA is heated at 65° for 2 min to melt any secondary structure of the single-strand cDNA, which would decrease the tailing efficiency in some cases. After heating, transfer the sample to ice to prepare it for the reaction.

To the cDNA sample, in a final volume of 50 μl, add

10× TdT buffer [2 *M* potassium cacodylate (pH 7.2), 10 m*M* $MgCl_2$, 10 m*M* 2-mercaptoethanol]: 5 μl

dGTP (50 μM): 5 μl, to a final concentration of 5 μM

$CoCl_2$ (10 m*M*): 5 μl

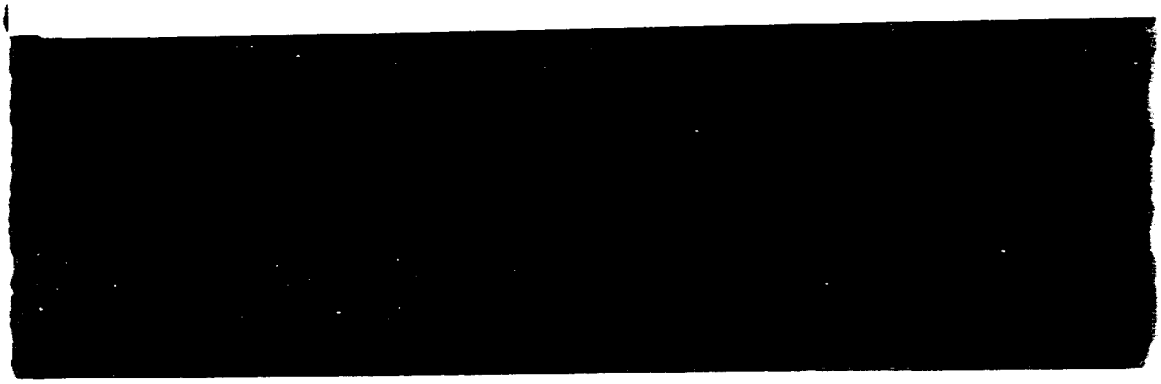
Terminal deoxynucleotidyltransferase: 40 units

Mix and transfer 3 μl to the control, tail-check tube, containing 0.5 μl of the labeled oligonucleotide. Incubate both cDNA and tail-check reaction for 30 min at 37°.

Stopping Tailing Reaction and cDNA Extraction–Precipitation. At the end of the tailing reaction, stop the main reaction (47 μl) by adding 1 μl of 0.5 *M* EDTA to the sample. If the reaction is not controlled, the resulting tail may be excessively long, causing difficulties in sequencing. Perform the proteinase K digestion and phenol–chloroform and chloroform extraction and back extraction as previously described for the first-strand cDNA reaction, and precipitate with 2.5 vol of ethanol after adding 0.2 *M* NaCl, as previously described. Further addition of carrier is not necessary because the glycogen and tRNA are still present from the previous steps. Finally, resuspend the pellet in 39 μl of 0.1× TE. In a parallel procedure, load the check-gel as explained in the next section.

Tailing Reaction Check. Use a small aliquot of the reaction to check the tailing efficiency. We usually prepare gels with 10% acrylamide, 8 *M* urea, and 1× TBE, and of size 400 × 150–200 × 0.25 mm (thickness). One side of the glass is treated with Sigmacote (Sigma) for easy removal of the gel after the electrophoresis. Thirty-five milliliters of acrylamide is polymerized by adding 35 μl of *N,N,N′,N′*-tetramethylethylenediamine (TEMED) and 70 μl of 25% ammonium persulfate (APS). For more details about the acrylamide gel electrophoresis, refer to Sambrook *et al.*[9]

Load the sample after adding 1 vol (3.5 μl) of sequence loading buffer



ıcleotide kinase (PNK), following
nol–chloroform and chloroform,
g of glycogen as a carrier.[9] Dilute
pproximately 1 ng/μl. The labeled
ınd can be used for no more than
ıis oligonucleotide, the sequence
ɛ are no strong internal secondary
·ngth of the G tail in a parallel re-

ted at 65° for 2 min to melt any
DNA, which would decrease the
·ating, transfer the sample to ice

ne of 50 μl, add
ɔdylate (pH 7.2), 10 m*M* $MgCl_2$,

:ntration of 5 μM

ɛ: 40 units
ail-check tube, containing 0.5 μl
ɔth cDNA and tail-check reaction

Extraction–Precipitation. At the
reaction (47 μl) by adding 1 μl
on is not controlled, the resulting
ulties in sequencing. Perform the
ɔform and chloroform extraction
ibed for the first-strand cDNA
thanol after adding 0.2 *M* NaCl,
·f carrier is not necessary because
rom the previous steps. Finally,
In a parallel procedure, load the

ıliquot of the reaction to check
gels with 10% acrylamide, 8 *M*
·200 × 0.25 mm (thickness). One
·te (Sigma) for easy removal of
·five milliliters of acrylamide is
,*N*′-tetramethylethylenediamine
·rsulfate (APS). For more details
refer to Sambrook *et al.*[9]
5 μl) of sequence loading buffer

(95% deionized formamide, 0.05% bromophenol blue, 0.05% xylene cyanol, 10 m*M* EDTA) and denature at 65° for 3 min. Usually half of the control tailing reaction (3.5-μl aliquot) mixture is loaded with an equivalent amount of nontailed, kinased starting oligonucleotide (0.5 μl plus 3 μl of water plus 3.5 μl of loading buffer). A marker such as P-Bluescript, *Hpa*II cleaved and radiolabeled, can also be loaded using the same loading buffer. Convenient labeling can be done by filling in the *Hpa*II site with Klenow fragment in the restriction buffer.[9] The gel can be run at constant power (35 W) until the bromphenol dye approximates the bottom of the gel; usually 60 min of running is sufficient. Finally, dry the gel for 20–30 min at 80° with a gel dryer. The exposure time is as short as 30 min for an image analyzer [for instance, using a Fuji (Tokyo, Japan) BAS 2000 image analyzer] or a few hours if conventional films are used.[9]

By measuring the difference in length of the tail-check reaction and the starting oligonucleotide, the efficiency and length of tailing can be estimated. A satisfactory tailing reaction shows a shift in mobility of the oligonucleotide of approximately 15 bases (± two or three), with a narrow distribution of the G tail. G-stretches of such a length do not usually interfere with the sequencing operations.

cDNA Preparation: Second-Strand Synthesis

As a primer for the second-strand cDNA, an oligonucleotide of sequence 5′-$(GA)_9$(GAGCTCACTAGTC$_{11}$-3′ and containing the *Sst*I site is prepared and purified by commonly used techniques, for instance, as for the first-strand cDNA primer. To optimize the priming of the second-strand cDNA just at the end of the oligo(dG) stretch and to reconstitute the restriction sites, a temperature gradient is employed for annealing. In fact, by using a temperature ramp, mismatches in the oligo(dG) or oligo(dC) can be destabilized to obtain perfectly annealed primer/dG tails.[4] Hot-start priming is thus performed after setting the second-strand program on the thermal cycler as follows:

Step 1: 5 min at 55°
Step 2: Negative ramp of −20° (until 35°), with slope of −0.3°/1 min
Step 3: 35° for 10 min
Step 4: 68° for 20 min; repeat from step 3 twice (total, three times)
Step 5: +4°

Prepare tubes A and B containing:
Tube A:
Oligo(dG)-tailed cDNA: 39 μl
Second-strand *Sst*I primer adapter (100 ng/μl): 6 μl

Second-strand buffer [200 m*M* Tris (pH 8.93), 350 m*M* KCl, 50 m*M* $(NH_4)_2SO_4$, 10 m*M* $MgSO_4$, 10 m*M* $MgCl_2$, 0.5% Triton X-100, BSA (0.5 mg/ml]: 6 μl
dNTPs (2.5 m*M* each): 6 μl

Tube B:

[α-^{32}P]dGTP: 0.5 μl

After starting the second-strand program, put tubes A and B on the thermal cycler. Add to tube A 3 μl of Ex-*Taq* polymerase (5 U/μl) when the samples are at 55°, during the first step. Mix quickly but thoroughly, and immediately transfer a 5-μl aliquot to a new tube containing 0.5 μl of [α-^{32}P]dGTP. The transfer should be immediate to avoid the start of polymerization before mixing with the radioisotope-labeled nucleotide, which may lead to underestimation of the final cDNA yield. At this stage, any [α-^{32}P]dNTP can be used instead of [α-^{32}P]dGTP.

Spot a 0.5-μl aliquot from the cold 55-μl reaction sample and the hot 5-μl reaction sample onto DEAE paper (DE-81). Count the radioactivity before and after the washing with 0.5 *M* sodium phosphate and calculate the yield of the second strand after subtracting the contribution of the first stand (as obtained from the 0.5-μl spot from the 55-μl cold reaction sample). Because 20.4 μg of dNTP, calculated as for the first-strand reaction, is used in the above reaction, the yield of the second-strand cDNA is obtained by

$$\text{Second strand } (\mu g) = 20.4 \times (\% \text{ of incorporation})/100$$

Typical yields may vary between 100 and 500 ng, but even if the yield is as low as 50 ng there is still enough cDNA to make a representative library. The total double-strand cDNA quantity is determined simply by multiplying by 2 the quantity of second-strand cDNA. Keep a record of the counts per minute obtained at this point to use in a later stage to calculate the units of enzyme to restrict the cDNA and the ligation conditions.

A 0.5-μl aliquot of the second strand may be run on an alkaline gel to check the size after the second strand. It should approximately reflect the size of the first-strand cDNA and may help to estimate the average size to calculate the vector-to-insert ratio for the ligation (see below). The remaining aliquot of the radioactive reaction and the 55-μl main reaction can be combined if the CPM contribution of the first strand (counts per minute) seems to be too low to be followed in the later purification stages (less than 5000–10,000 cpm), or if there is some reason to suspect that the yield will be low.

The cDNA is treated as described above with proteinase K, organic extraction and back extraction, precipitation with ethanol using NaCl as a salt, and washing with 70% ethanol. Finally, resuspend the cDNA in 45 μl of 0.1× TE.

Cleaving cDNA. Calculate the units of restriction enzymes to be used to cleave the cDNA, based on the yield of the second-strand synthesis and assuming that no consistent fraction is lost during the precipitation. cDNA is restricted by using 25 U/μg of both *Sst*I and *Xho*I. These units have been tested and do not cleave the hemimethylated sites of the cDNA.[8] The restriction enzymes must be of high quality to prevent problems with overdigestion.

To the redissolved cDNA add, in a final volume of 50 μl:

Restriction buffer M (10×; BRL): 5 μl

25 U of *Sst*I per μg of double-stranded cDNA; in most cases, this quantity will be scaled down (for instance, for 200 ng of double-stranded cDNA use 5 U of enzyme)

Perform the following steps:

1. Incubate for 1 hr at 37°
2. Inactivate the restriction enzyme *Sst*I by incubating at 65° for 10 min.
3. Cool the sample on ice for 2 min and add additional NaCl to a final colume of 100 m*M* (1 μl of 2.5 *M* NaCl) to prepare high buffer conditions for *Xho*I.
4. Add the same amount (units) of *Xho*I as previously used for *Sst*I.
5. Incubate for 1 hr at 37°.

During the reaction, equilibrate a CL-4B column as previously for the purification of cDNA before tailing.

Stop the reaction by treating as described above with proteinase K, then conduct phenol–chloroform extraction once and back extraction. Chloroform extraction is not necessary at this stage. Keep the back-extracted aliquot separated until the CL-4B column is equilibrated. Apply the main fraction of cDNA, followed by the back-extracted aliquot, to the CL-4B column and centrifuge for 2 min at 400*g*. Collect the eluted cDNA in a new tube and apply 50 μl of the column buffer; centrifuge again at 400*g* for 2 min. Collect the fraction and transfer to a separate tube. Repeat this step two or three more times, keeping the fractions separate at all times.

Count by scintillation the tubes containing the cDNA. In most cases, the bulk of the cDNA is contained in the first two or three fractions. Subsequent fractions, which usually contribute only a small part of the cDNA (but that may contain first-strand/second-strand cDNA primer dimers), can be discarded. When precautions were not taken, we observed contamination by the dimer of the oligo(dG)-tailed first-strand cDNA primer and second primer. By selecting only the initial fractions, this risk can be greatly minimized.

Finally, precipitate with ethanol the selected, pooled fractions, as previously described, and wash with 70% ethanol. Before resuspending the

cDNA, count the pellet by scintillation. This will allow calculation of the cDNA recovery after the CL-4B column and ethanol precipitation, for the next ligation step.

Ligation of cDNA into λ DNA Purified Arms. Roughly estimate the average size of the cDNA from the second-strand alkaline gel. If the size of the second-strand cDNA has not been checked by electrophoresis, assume that it has an average size of 2 kb when calculating the optimal ratio between cDNA and vector. A high vector-to-cDNA ratio will increase the background but decrease the formation of chimeric clones. Consequently, a high vector-to-cDNA (2:1) ratio should be used only in the case of a very low background vector. Notice that an increase in the vector concentration does not increase the yield of cloning, because for the packaging reaction the concatameric form of λ DNA has a higher efficiency than do single arms ligated with an insert. The highest efficiency of packaging is thus obtained with a 1:1 ratio between the vector and insert.

If the background of the vector is not very low (higher than 2.5%), the ratio of vector to cDNA should be between 1:0.8 and 1:1. When the vector-to-cDNA ratio is higher than 1:2 (to overcome the background of the vector), the contamination of chimeric clones is usually not acceptable. Alternatively, small aliquots of cDNA can be ligated at different cDNA-to-vector ratios (for instance, 1:3, 1:1, and 3:1) to search for the best compromise between the efficiency, background, and possibility of formation of chimeric clones.

Set up the ligations in a final volume of 5 μl (use up to 5 μg of vector), with the amount of cDNA selected. Ligate overnight at 16°, and subsequently conduct the *in vitro* packaging. An additional control experiment on the vector background can be done by ligating an aliquot of the λ vector without the cDNA. Usually the *in vitro* packaging reaction is done with at least half (25 μl) of the packaging extract Max Plax when using up to 5 μg of λ vector. Standard protocols are followed, as suggested by the manufacturer, for all of the subsequent manipulations.[9] Usually, the entire procedure can be expected to yield at least 10^6–10^7 primary plaque-forming units (PFU).

The library is then ready for the next analysis or amplification.[9]

To test the quality of the library, check the size of 20–30 clones by one of several methods: (1) direct PCR of the lysate by long PCR[6]; (2) λ miniprep followed by restriction digestion[21]; (3) *in vivo* excision, followed by checking the size of the excised plasmid by restriction digestion.[10,11]

The average size of a satisfactory library is 1.5 kb, and 2 kb in the case of an excellent library.

Additional controls to ensure the quality of the library may include random sequencing of 50 clones from the 5′ end (*Sst*I side). The sequences

This will allow calculation of the
ınd ethanol precipitation, for the

ied Arms. Roughly estimate the
ıd-strand alkaline gel. If the size
ı checked by electrophoresis, as-
hen calculating the optimal ratio
r-to-cDNA ratio will increase the
›f chimeric clones. Consequently,
ld be used only in the case of a
ı increase in the vector concentra-
ıing, because for the packaging
A has a higher efficiency than do
ıighest efficiency of packaging is
he vector and insert.
very low (higher than 2.5%), the
ween 1:0.8 and 1:1. When the
(to overcome the background of
c clones is usually not acceptable.
ın be ligated at different cDNA-
and 3:1) to search for the best
ground, and possibility of forma-

›f 5 μl (use up to 5 μg of vector),
:ate overnight at 16°, and subse-
An additional control experiment
ligating an aliquot of the λ vector
'ackaging reaction is done with at
Max Plax when using up to 5 μg
'ed, as suggested by the manufac-
ıns.[9] Usually, the entire procedure
0^7 primary plaque-forming units

. analysis or amplification.[9]
:k the size of 20–30 clones by one
the lysate by long PCR[6]; (2) λ
n[21]; (3) *in vivo* excision, followed
ıid by restriction digestion.[10,11]
ary is 1.5 kb, and 2 kb in the case

ıality of the library may include
5′ end (*Sst*I side). The sequences

of several clones will be present in GenBank and can be checked if they are full length: more than 90–95% of identified clones should be full-length.

Alternatively, to quickly check the full-length content, two replica filters containing 50.000 PFU can be hybridized with probes 3′ and 5′ to a given cDNA, as full-length clones will hybridize to both. For instance, EF-1α (1.7 kb) and GAPDH (1.2 kb) are ubiquitously and highly expressed and

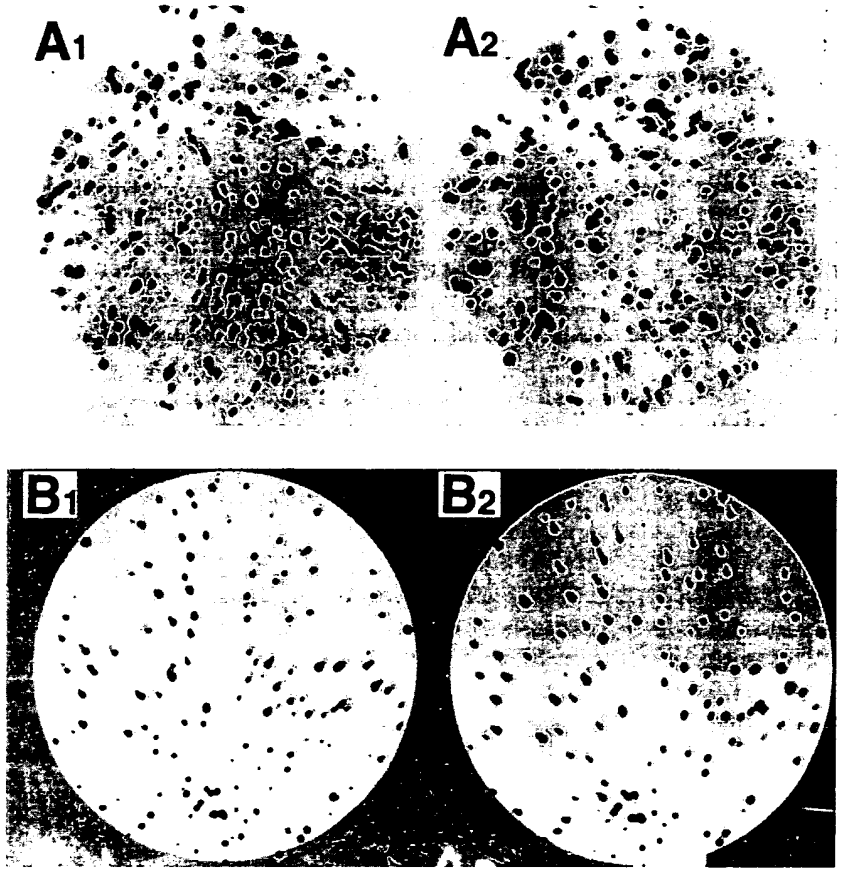


FIG. 2. Assessment of full-length cDNA content of a library by hybridization. A1 and A2, Hybridization with probes, respectively, at the 5′ end (from nucleotide 27 to 455) and 3′ end (from nucleotide 880 to 1175). About 98% of clones positive for the 3′-end probe are also positive for the 5′-end probe. B1 and B2, Hybridization with probes, respectively, at the 5′ end (from nucleotide 34 to 365) and 3′ end (from nucleotide 1264 to 1643) of EF-1α. More than 95% of the clones positive for the 3′ probe are also positive for the 5′ probe.

give, respectively, more than 95 and 98% of plaques positive for both 3′ and 5′ probes, in the case of a satisfactory full-length library (Fig. 2).[4]

Alternative Cloning Strategies

Our methodology can in principle be applied to a wide range of cloning vectors, including different λ vectors and plasmids, provided that they contain restriction sites that are sensitive to hemimethylation. When *Sst*I and/or *Xho*I need to substituted, different restriction sites can be designed on the primer adapters that are suitable for other vectors. We tested the resistance of hemimethylated cDNA to restriction digestion at a high enzyme-to-substrate ratio and identified several enzymes that are suitable for such a strategy and are present in many multiple cloning sites, such as *Bam*HI, *Not*I, *Apa*I, *Acc*I, *Bst*XI, *Eag*I, and *Sal*I. However, we do not recommend using *Not*I or *Eag*I next to the dC stretch of the second-strand primer because of subsequent difficulties in sequencing.[8] See also Nelson *et al.*[7] for an extensive list of enzymes, although we found, in contrast to what was described, that *Sac*I can partially cleave hemimethylated cDNA.[8]

Our method, using the powerful thermostabilized RT and the biotinylated cap trapper, makes possible efficient preparation of full-length cDNAs longer than 10 kb. The availability of full-length cDNA libraries can greatly facilitate cloning work, particularly in large-scale sequencing projects.

Acknowledgments

This study was supported by Special Coordination Funds and a Research Grant for the Genome Exploration Research Project from the Science and Technology Agency of the Japanese Government, and by a Grant-in-Aid for Scientific Research on Priority Areas and the Human Genome Program, from the Ministry of Education and Culture, Japan, to Y.H. This work was also supported by funding from the Core Research of the Evolutional Science and Technology Program from JST. Support also came from a Grant for Research on Aging and Health, and a Grant-in-Aid for a Second Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare of the Japanese Government to Y.H.

STIC-ILL

@P606.D46P358
APL ERROR

Fr m: Lacourciere, Karen
Sent: Wednesday, July 24, 2002 10:31 AM
To: STIC-ILL
Subject: ill order

Please provide the following reference for use in examining 09/836,410

FASEB Journal, (March 12, 1999) Vol. 13, No. 4 PART 1, pp. A531.

IOVS, (March 15, 2001) Vol. 42, No. 4, pp. S94. print.

INVESTIGATIVE OPHTHALMOLOGY AND VISUAL SCIENCE, (2001 Nov) 42 (12) 3000-7.

Meth. Enzymol. 303, 19-44, (1999).

Genome Res. 10 (10) 1617-1630 (2000)

Genome res. 10 (11) 1757-1771 (2000)

Thank-You!

Karen A. Lacourciere Ph.D.
CM1 11D09 GAU 1635
(703) 308-7523

Normalization and Subtraction of Cap-Trapper-Selected cDNAs to Prepare Full-Length cDNA Libraries for Rapid Discovery of New Genes

Piero Carninci,[1] Yuko Shibata, Norihito Hayatsu, Yuichi Sugahara, Kazuhiro Shibata, Masayoshi Itoh, Hideaki Konno, Yasushi Okazaki, Masami Muramatsu, and Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC) and Genome Science Laboratory, RIKEN Tsukuba Institute, Core Research of Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), Tsukuba 305-0074, Japan

In the effort to prepare the mouse full-length cDNA encyclopedia, we previously developed several techniques to prepare and select full-length cDNAs. To increase the number of different cDNAs, we introduce here a strategy to prepare normalized and subtracted cDNA libraries in a single step. The method is based on hybridization of the first-strand, full-length cDNA with several RNA drivers, including starting mRNA as the normalizing driver and run-off transcripts from minilibraries containing highly expressed genes, rearrayed clones, and previously sequenced cDNAs as subtracting drivers. Our method keeps the proportion of full-length cDNAs in the subtracted/normalized library high. Moreover, our method dramatically enhances the discovery of new genes as compared to results obtained by using standard, full-length cDNA libraries. This procedure can be extended to the preparation of full-length cDNA encyclopedias from other organisms.

It has been tempting to prepare and use full-length cDNA libraries (Kato et al. 1994; Maruyama and Sugano 1994; Edery et al. 1995; Carninci at al. 1996, 1998; Carninci and Hayashizaki 1999) in large-scale gene discovery efforts incorporating one-pass sequencing that resemble the existing EST projects (Adams et al. 1991, 1995; Hillier et al. 1996; Marra et al. 1999). One advantage of such an approach is that most clones contain the complete coding sequence as well as the 5′ and 3′ untranslated regions (UTRs), thus dramatically accelerating the subsequent sequencing, biocomputation, and protein expression and other functional assays. However, generating full-length cDNA libraries has some inherent problems. The preparation of full-length cDNA is more efficient for short mRNAs than for long transcripts. In addition, cloning and propagation is more difficult for long cDNAs than short cDNAs, thus introducing further size bias. Using truncated cDNAs to retrieve the full-length cognate is impractical on the genomic scale; however, cDNAs in a standard library can be cloned in either their full-length or truncated forms, thus favoring discovery of at least one EST for any gene, regardless of its length.

Another problem associated with gene discovery reflects the nature of the cellular mRNA. Depending on their expression, mRNAs can be defined as superprevalent (or abundant), intermediate, or rare. In a typical cell, 5–10 species of superprevalent cDNA comprise at least 20% of the mass of mRNA, 500–2000 species of intermediately expressed mRNA comprise 40%–60% of the mRNA mass, and 10,000–20,000 rare messages may account for <20%–40% of the mRNA mass. This average distribution may vary markedly between tissue sources, and the presence of numerous highly expressed genes may further unbalance this distribution. Sequencing cDNAs from standard cDNA libraries is ineffective for discovering rarely expressed genes, when intermediately and highly expressed cDNAs would be sequenced redundantly.

We are working on the mouse full-length cDNA encyclopedia project, the ultimate goal of which is to collect at least one full-length cDNA for every expressed gene, regardless of the tissue (http://genome.rtc.riken.go.jp/). To this end, we wanted to remove not only redundant cDNAs but also sequences that were represented already in a previous library, thus accelerating the discovery of new, full-length cDNAs. Therefore, we wanted to develop a technology capable not only of normalizing the frequencies of full-length cDNAs from mRNAs belonging to the three different classes of expression but also of subtracting cDNAs that have already appeared in other libraries.

[1]**Corresponding author.**
E-MAIL rgscerg@rtc.riken.go.jp or carninci@rtc.riken.go.jp; FAX 81-298-369098.
Article and publication are at www.genome.org/cgi/doi/10.1101/gr.145100.

We considered several possible strategies that were based on the reassociation kinetics of nucleic acids, but none were amenable to full-length cDNA approaches. Existing technologies (Soares et al. 1994; Bonaldo et al. 1996) widely used for normalization and subtraction for large-scale gene discovery through EST approaches were unappealing to us, mainly because they were not easily applicable to long cDNA inserts. These protocols rely in fact on the reassociation of the nucleic acids in amplified plasmid libraries. However, plasmid libraries are associated with a cDNA-size cloning bias that manifests as an increased cloning efficiency of short cDNAs. In addition, during library amplification before normalization and/or subtraction, the growth of cDNA clones varies with plasmid length; therefore, long clones are underrepresented after bulk amplification of the library. This discrepancy would lead to underrepresentation of long cDNAs and difficulty in cloning long, rare cDNAs.

To avoid the problems related to amplification of libraries, we wanted to develop a technique to normalize and subtract cDNA before cloning. Published protocols did not lead to equal representation among clones of different sizes, maintain the length of long cDNAs after hybridization, or incorporate simultaneous normalization and subtraction of cDNAs. Therefore, methods based on PCR (Takahashi and Ko 1994; Diatchenko et al. 1996) in which long and otherwise difficult-to-amplify cDNAs are likely to be underrepresented were unsuitable for a full-length cDNA approach. Methods in which an immobilized nucleic cDNA driver on a solid matrix to subtract mRNA tester (Sasaki et al. 1994; Tanaka et al. 1996) were unsuitable for our purposes because of the risk of mRNA degradation before cDNA synthesis. In addition, the hybridization kinetics of nucleic acids immobilized on a solid phase (Tanaka et al. 1996) is slower than those for solution hybridization (Anderson and Young 1985). Libraries created with PCR- and solid matrix–based technologies were only partially characterized and showed sequence redundancy similar to that of nonnormalized cDNA libraries used in ESTs projects.

In addressing the normalization of full-length cDNAs, we felt that an aliquot of the mRNA initially used for the cDNA library preparation would be the ideal driver because it reflects the complexity of the first-strand cDNA tester. In addition, such a strategy could be extended easily to subtract sense mRNA sources from other tissues. Further, because cDNA cloning vectors commonly used in cDNA libraries construction carry the promoter sequences of T7 and T3 RNA polymerases, it would be easy to subtract cDNAs obtained from other libraries or pools of clones that had been already categorized by one-pass sequencing. Although frequently used to separate the hybridized driver and tester, hydroxyapatite chromatography requires strict temperature control, thus rendering the procedure technically demanding. Biotinylation of the mRNA driver is an easy alternative that is amenable to upscaling. Further, biotinylation can be coupled easily to streptavidin-phenol extraction (Barr and Emanuel 1990) or techniques using magnetic beads, provided that the reported cDNA degradation caused by photobiotinylated drivers (Fargnoli et al. 1990) is prevented. Here we present the first method for preparing normalized/subtracted libraries that also facilitates high-efficiency cloning of full-length cDNAs.

RESULTS

Strategy

In preliminary experiments, we first aimed at developing or adapting technologies that addressed the following points: high-efficiency removal of mRNA drivers; lack of cDNA size reduction after hybridization that would affect the frequency of full-length cDNAs; suitability for both normalization and subtraction; low cross-reactivity between similar but unidentical sequences; and being reproducible and amenable to upscaling both in terms of size of the driver and the number of libraries to be prepared.

Our general strategy (Fig. 1) involving hybridization of first-strand cDNA to mRNA has several advantages. This methodology is a modification of the previously proposed cDNA library preparation by Cap-Trapper (Carninci and Hayashizaki 1999) and accommodates the cloning of full-length, normalized-subtracted cDNA. Our method has the benefit of being amenable to using starting mRNA for the normalization process as well as to subtraction with an in vitro–transcribed RNA driver from any other directionally cloned cDNA library, preferably one made by using Cap-Trapper technology. After the subtraction/normalization step, cDNA is cloned.

Development of the Technology

We incorporated hybridization in formamide at 42°C because the mild temperature apparently did not lead to degradation of the cDNA after prolonged incubation (not shown). We checked whether relevant nonspecific hybridization occurred at these conditions to avoid removal of related but different sequences. We used two clones that share 76.8% identity in 1554 nucleotides, which also had long stretches of ~85% homology. These clones were the mouse full-length tubulin-M β5 and an unknown mouse cDNA that is 93% similar to the Chinese hamster mRNA for beta tubulin (clone B3T). Hybridization in 0.25 M NaCl gave excellent removal of specifically hybridized clones without cross-hybridization between the two clones.

One of our primary requirements was that our

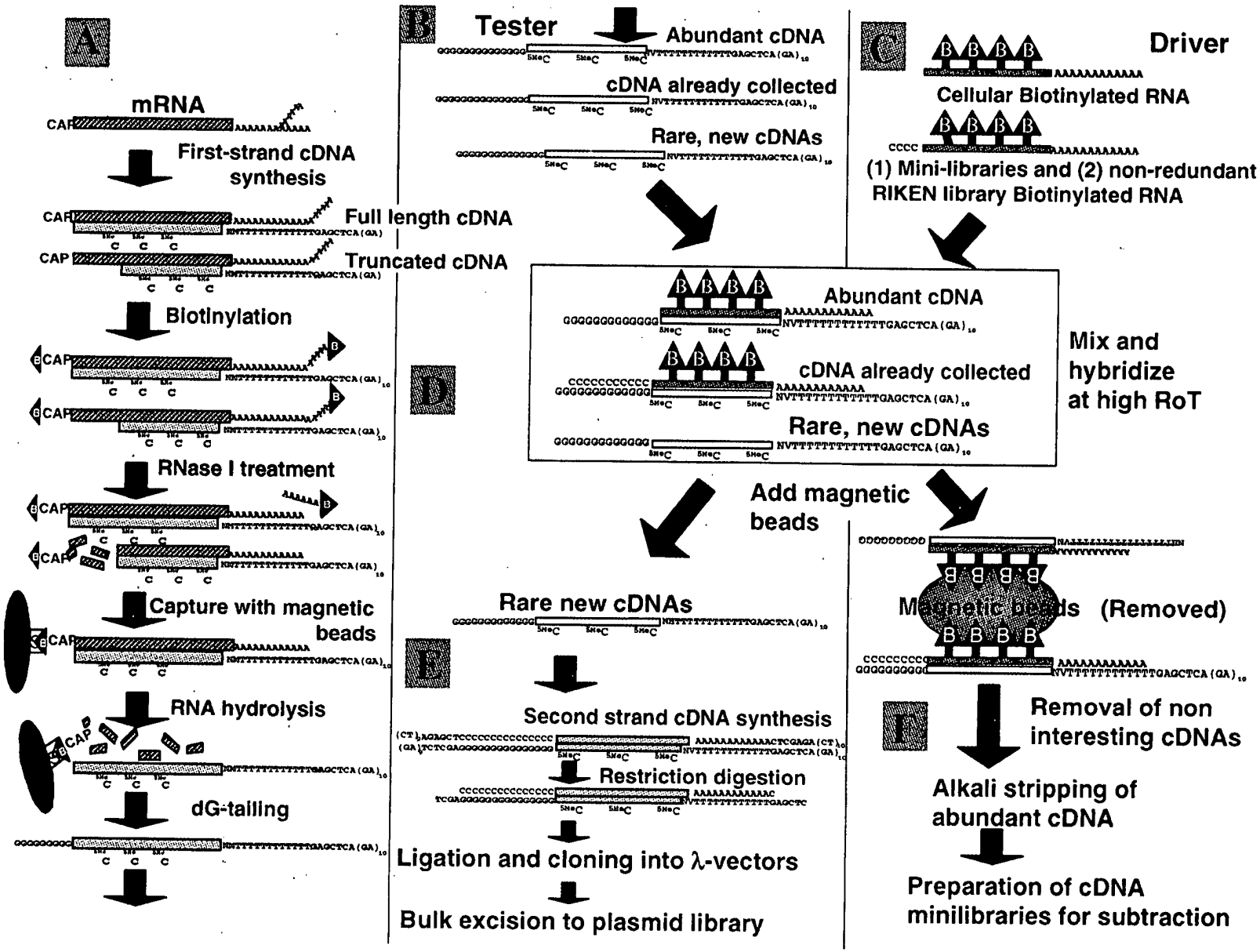


Figure 1 Schematic diagram of the normalized-subtracted cDNA preparation protocol. (*A*) general scheme for preparing full-length single-strand cDNA; (*B*) representation of various populations of tester cDNAs; (*C*) normalizing driver (cellular mRNA) and subtracting drivers (run-off transcripts); (*D*) hybridization; (*E*) rare/new cDNAs are used for second-strand cDNA preparation (normalized/subtracted cDNA library); (*F*) abundant cDNAs/unwanted cDNAs are removed and may be used for the preparation of minilibraries to implement subtraction.

method had to remove most of the RNA driver, a condition essential for removal of the tester-driver hybrid as well. Therefore, we tested whether the RNA biotinylation system afforded a high efficiency of labeling RNAs with biotin. The method that initially worked most efficiently to remove the driver was psoralen-biotinylation, which enabled removal of most of the biotinylated driver by using streptavidin beads. We verified the utility of our selected methods by hybridizing a 5-Kb tester cDNA, at the 3'-end of reeler cDNA (Hirotsune et al. 1995), to its RNA driver at RoT = 0.2 (for a detailed description of RoT see Anderson and Young 1985). This methodology led to removal of ~98%–99% of the starting cDNA, as measured by counting the radioactivity of the cDNA and by visualizing the intensity of electrophoresis smear. Because it performed as well as the psoralen-biotin system but was easier to use, we switched to the Mirus biotinylation kit (Panvera) for subsequent library preparation. To subtract the hybridized probe, magnetic porous glass (MPG) streptavidin beads (CPG) worked better in our hands than did other systems, such as the streptavidin-phenol technique; the streptavidin beads removed close to or >99% of tester-driver hybrid (not shown).

Reduction of the Frequency of Highly Expressed cDNAs

After preliminary experiments, we prepared several normalized and normalized/subtracted libraries (Table 1). cDNA libraries that were generated more recently were normalized and subtracted with both the minilibraries and the RNA drivers derived from the rearrayed nonredundant RIKEN cDNA encyclopedia to reduce wasteful resequencing of clones already represented.

We compared the second-strand cDNA from a standard pancreas cDNA library to its normalized/subtracted counterpart (Fig. 2). Normalization was performed at RoT = 10, and subtraction was accomplished by using a set of minilibraries, each of which contained 1000–2000 redundant, mainly abundant clones from liver, lung, brain, or placenta. We generated the minilibraries by cloning the highly expressed fraction of previously prepared, normalized cDNA libraries. Amplified cDNA minilibraries were then used to prepare the subtracting drivers (see Methods). The RoT for the subtracting drivers equaled 1 U for every 200 clones (e.g., RoT = 5 when 1000 clones were used). The average size of normalized, subtracted cDNA was longer than that of the nonnormalized, nonsubtracted cDNA, suggesting that degradation does not occur during the subtraction step and that long cDNAs are expressed more rarely than the short ones. We frequently have observed similar results with other cDNAs. In addition, the cDNAs corresponding to highly expressed cDNAs are absent from the normalized-subtracted library; the electrophoretic pattern demonstrates the cDNA normalization (Fig. 2). Working with full-length cDNA further helped to better visualize the removal of specific cDNAs. We did not further sequence the standard pancreas cDNA library because of the superprevalence of a very few cDNA species. Another way of demonstrating the benefit of normalization/subtraction is shown in Figure 3. We used first-strand cDNA from lung as a template and normalized or normalized/subtracted one aliquot and cloned another aliquot that was not normalized. Plaque hybridization of the normalized cDNA library and the standard counterpart suggested that the frequency of highly expressed genes was reduced in the normalized library. When we screened 10,000 plaques of the normalized lung library, the representation of elongation factor 1-α was reduced from 90 plaques in the control library to 10 in the normalized library, carbonyl reductase decreased from approximately 70 to 3, and uteroglobin was reduced from approximately 510 to 2 plaques. These results strongly suggest that the frequency of highly expressed cDNAs in the normalized library was much less than that in the control.

Increasing the Frequency of Discovering Rare Genes

To verify the enrichment of rare cDNAs, large-scale sequencing of the library is the most indicative test because we anticipated reduced sequence redundancy and increased discovery of new genes. We prepared several libraries (Table 1) and assessed them by checking the titer, average size of the cDNA inserts, presence of full-length cDNA, redundancy (by sequencing the 3'-ends of inserts), and recovery of new genes/ESTs. Assessing the degree of sequence redundancy was our final evaluation of the efficiency of the normalization/subtraction process. Standard libraries (series 22–, 23–, 26–, and 31–) prepared from an aliquot of the starting cDNA are shown for comparison (Table 1).

From Lib.32, we developed a new cloning vector that could incorporate long cDNA inserts, thereby increasing the efficiency of cloning for long cDNAs and facilitating bulk excision into a plasmid cDNA library by using the *cre-lox* system (P. Carninci, in prep.). Normalized/subtracted libraries incorporating this cloning system are deeper than previously prepared libraries, in which mainly short cDNAs were cloned. In a successful normalized-subtracted cDNA library (e.g., Lib.49 from testicular tissue) using our new cloning system, the redundancy of the sequences from 3'-ends was as low as 1.63 (calculated by dividing the number of different clusters by the total number of sequencing passes) after sequencing 8900 clones. Redundancies of <2.0 in >10,000–15,000 3'-end sequences can be expected in successful cDNA libraries from complex tissues (e.g., testis, brain, and thymus). Further, the normalized/subtracted cDNA libraries facilitated efficient and in-

creased recovery of unknown genes. For example, Libraries 22–100, 23–100, 26–100, and 31–100 produced about or more than twice the amount of new data per sequencing reaction than did the standard library counterparts 22–000, 23–000, 26–000, and 31–000 (Table 1, no-EST and no-NT columns).

Sequencing several cDNAs from various libraries reveals a relevant decrease in sequence redundancy in the normalized-subtracted library as compared to that in standard cDNA libraries (Fig. 4). Normalization increases the frequency of new gene discovery to almost twice that for standard libraries during a given sequencing effort. In comparison, subtraction with nonredundant, rearrayed drivers removes cDNAs redundant among various tissues and therefore improves the rate of new gene discovery during the course of the project.

To date, by sequencing 929,814 clones, we have been able to cluster 128,671 3′-end sequences into different groups. Because of the constant monitoring of the gene discovery rate per given cDNA library (P. Carninci, submitted), normalized/subtracted cDNA libraries were largely preferred over standard counterparts. In addition, 60,941 singletons (clusters of clones that appeared only once) were collected from 829,017 sequencing runs from normalized/subtracted cDNA libraries. Currently, we have rearrayed approximately 30,000 cDNA clones to be used for preparing RNA drivers for new cDNA libraries for the mouse cDNA encyclopedia project (http://genome.rtc.riken.go.jp/). Thus far we have prepared normalized, subtracted cDNA libraries at RoTs >200 that were producing 20%–30% of new sequences against our internal database by 3′-end reading when comparing against 70,000 different cDNA clusters.

Full-Length cDNA Rate

Of primary importance is that the full-length cDNA content is maximal after the normalization/subtraction steps. In fact, in Table 1, we can appreciate the proportion of full-length cDNAs in the various cDNA libraries. The evaluation was performed as summarized (Y. Sugahara et al., submitted). We then sequenced several hundreds of clones from the normalized-subtracted libraries. Sequences that hit "complete mRNA" sequences of mouse were aligned and checked for the presence of the initiator ATG. The presence of the initiator ATG was the factor used to assess the quality of 5′-ends instead of the exact overlap of our clones with published 5′ sequences. In fact, published "complete" sequences may differ from Cap-Trapper sequences because of differences in promoter/transcriptional start-site usage and cloning techniques. In contrast, the presence of the initiator ATG reliably shows that a given clone is practically full-length. In most of our cDNA libraries (Table 1), 80%–100% of clones included the first ATG, with an average of 88.1% in the libraries here presented. This average value goes close to standard Cap-Trapper cDNA libraries, where about 95% of full-coding cDNA was reported (Carninci et al. 1996), although the data set used for this previous analysis was different because we included the comparison of mouse to homologue genes of other vertebrates. A successful blastocyst cDNA library obtained with a cap-switch method (Sasaki et al. 1998) scored similarly (94% of clones contained the first ATG), but in this library we could cluster only 937 genes in 3995 sequencing passes. In data of another project (Marra et al. 1999), in three nonnormalized, full-length oligo-capping cDNA libraries (Maruyama and Sugano 1994), about 77% of clones contained the first ATG (Sugahara et al., submitted). By considering ESTs candidates as full-length at 5′ end when they match within 50 bp from a sequence annotated as full-length, these oligo-capping libraries were scored 65%–70% full-length at 5′ end (Marra et al. 1999), while ESTs from remaining standard and mainly normalized libraries (Bonaldo et al. 1996) scored about 27% full-length rate at 5′ end (Marra et al. 1999). Unfortunately, gene diversity from oligo-capping libraries was reduced: 2159/8231, 4463/21,594, and 2648/18,792 of "clusters versus classified ESTs" were obtained, respectively for the "mewa,", "mkia," and "mlia" ibraries as clustered in the Unigene database (http://www.ncbi.nlm.nih.gov/UniGene/Mm.Home.html) on July 18, 2000.

Specificity of the Normalization-Subtraction Steps

To preliminarily evaluate the specificity of our normalization-subtraction protocol, we checked the 3′ sequences of our libraries for the presence of B1 repeats. B1 repeats are present in about 5% of the 3′-ends of cDNAs. We assumed that if the hybridization were nonspecific, the frequency of B1 repeats, which are highly homologous, would be greatly reduced in normalized-subtracted libraries because of the excess of driver carrying the B1-repeat sequence. The frequency of B1 repeats apparently does not vary between normalized-subtracted and control libraries (Table 2), suggesting that the specificity of subtraction was satisfactory. The incidence of B1 regions differs from the previously described 5% because the sequencing read-length does not span the entire 3′ UTR. Detailed analysis of full-length cDNA sequences will confirm the specificity of subtraction among gene-family members.

DISCUSSION

When we began this work, genome-scale characterization of full-length cDNAs was an important problem. Although the technology for generating full-length cDNA libraries had been described already (Kato et al.

Table 1. Summary of Data on cDNA Libraries

Library ID	Developmental stage/tissue	Normalizing driver (Rot)	Subtracting driver (RoT)	Method	Titer	Insert size (kbp)	Sequencing	Species	Redundancy	No. ESTs (%)		No. NT (%)		Unique (%)		Coding (%)
18–100	Adult/pancrea	mRNA (5)	ms1 (20)	1	8.20e + 04	1.2	13556	3402	3.98	307	(9.0)	873	(25.7)	442	(13.0)	(100.0)
22–000		(standard)	(standard)		5.90e + 04	0.88	1458	488	2.99	26	(5.3)	52	(10.7)	42	(8.6)	(82.1)
22–100	Adult/ stomach	mRNA (5)	ms1 (20)	1	3.50e + 05	1.21	4400	1932	2.28	120	(6.2)	324	(16.8)	196	(10.1)	(82.1)
22–104		mRNA (5)	ms1 (20), Nm1 (5)	1	2.00e + 05	1.13	3936	1862	2.11	144	(7.7)	302	(16.2)	207	(11.1)	(82.1)
23–000		(standard)	(standard)		4.10e + 04		1179	556	2.12	30	(5.4)	50	(9.0)	36	(6.5)	76.8
23–100	Adult/ tongue	mRNA (5)	ms1 (20), Nm1 (5)	1	4.10e + 04	1.44	10267	4017	2.56	410	(10.2)	992	(24.7)	586	(14.6)	76.8
24–100	ES cell	mRNA (5)	ms1 (20), Nm1 (5)	1	1.30e + 05	1.77	15226	4495	3.39	236	(5.3)	677	(15.1)	485	(10.8)	(88.6)
25–100	Embryo13/ liver	mRNA (5)	ms1 (20), Nm1 (5)	1	8.50e + 04	1.19	5448	1525	3.57	52	(3.4)	179	(11.7)	168	(11.0)	(92.2)
26–000		(standard)	(standard)		6.10e + 05	1.38	2108	1061	1.99	31	(2.9)	97	(9.1)	71	(6.7)	92.3
26–100	Embry10/ whole boy	mRNA (7.5)	ms1 (30), Nm1 (7.5)	1	5.00e + 05	1.32	11267	4722	2.39	330	(7.0)	870	(18.4)	582	(12.3)	92.3
28–100	Embryo 10 + 11/ whole body	mRNA (7.5)	ms1 (30), Nm1 (7.5)	1	8.80e + 05	1.29	6248	3411	1.83	190	(5.6)	450	(13.2)	271	(7.9)	(93.9)
28–104		mRNA (7.5)	ms1 (30), Nm1 (7.5)	1	8.80e + 05	1.38	9321	4335	2.15	293	(6.8)	672	(15.5)	453	(10.4)	(93.9)
31–000		(standard)	(standard)		4.90e + 04	1.22	488	369	1.32	12	(3.3)	30	(8.1)	23	(6.2)	(86.2)
31–100	Embryo/ head	mRNA (10)	ms1 (40), Nm1 (10)	1	4.20e + 05	1.55	7838	4229	1.85	344	(8.1)	682	(16.1)	494	(11.7)	(86.2)
32–304	Embryo14 + 17/ head	mRNA (10)	ms1 (40), Nm1 (10) Nm2 (10)	1	3.30e + 05	2.5	424	389	1.09	22	(5.7)	41	(10.5)	20	(5.1)	(88.2)
38–304	Embryo11/ placenta & extraembryonic tissue	mRNA (10)	ms1 (40), Nm2 (10)	2	2.60e + 06	1.45	3657	2165	1.69	98	(4.5)	255	(11.8)	156	(7.2)	(100.0)
39–304	Embryo13/ whole body	mRNA (10)	ms1 (40), Nm1 (10) Nm2 (10)	2	2.10E + 05	2.47	348	319	1.09	14	(4.4)	33	(10.3)	22	(6.9)	(90.0)
49–304	Adult/testis	mRNA (10)	ms2 (90), Nm2 (10)	2	2.60E + 06	2.11	8900	5444	1.63	1102	(20.2)	1443	(26.5)	1214	(22.3)	(95.7)
52–304	Adult/ Xiphoid	total RNA (3)	ms2 (90), Nm2 (10)	2	7.30e + 05	2.69	272	256	1.06	12	(4.7)	21	(8.2)	15	(5.9)	(100.0)
53–304	Adult/ pituitary grand	total RNA (3)	ms2 (90), Nm2 (10)	2	2.10e + 06	2.38	8059	4658	1.73	411	(8.8)	640	(13.7)	833	(17.9)	(100.0)
54–304	Neonate6/ head	mRNA (10)	ms2 (90), Nm2 (10)	2	1.30e + 06	2.3	2663	2101	1.27	115	(5.5)	217	(10.3)	196	(9.3)	(90.0)
55–304	Neonate10/ head	mRNA (10)	ms2 (90), Nm2 (10)	2	1.70e + 06	2.18	603	525	1.15	39	(7.4)	83	(15.8)	44	(8.4)	77.3

Table 1. *(Continued)*

Library ID	Developmental stage/tissue	Subtracting driver (RoT)	Normalizing driver (RoT)	Method	Titer	Insert size (kbp)	Sequencing	Species	Redundancy	No. ESTs (%)	No. NT (%)	Unique (%)	Coding (%)
56–304	Embryo6/ whole body	ms2 (90) (subtracted only)	Nm2 (10)	2	6.00e + 05	2.3	416	371	1.12	10 (2.7)	25 (6.7)	16 (4.3)	(100.0)
57–304	Embryo8/ whole body	ms2 (90) (subtracted only)	Nm2 (10)	2	1.20e + 06	1.91	19632	7758	2.53	778 (10.0)	1598 (20.6)	1155 (14.9)	(100.0)
58–304	Adult/ thymus	ms2 (90) mRNA (10)	Nm2 (10)	2	1.70e + 06	3.27	10259	6442	1.59	604 (9.4)	1074 (16.7)	1100 (17.1)	(80.0)
60–304	Embryo13/ testis	ms2 (90) total RNA (5)	Nm2 (10)	2	5.70e + 05	1	1079	6498	1.7	672 (10.3)	1122 (17.3)	1243 (19.1)	(75.0)
61–304	Embryo14/ thymus	ms2 (90) (subtracted only)	Nm2 (10)	2	4.80e + 05	4.13	206	196	1.05	9 (4.6)	24 (12.2)	16 (8.2)	(60.0)
62–304	Embryo11/ head	ms2 (90) mRNA (10)	Nm2 (10)	2	3.30e + 05	2.19	2967	2374	1.25	149 (6.3)	265 (11.2)	256 (10.8)	(70.0)

The library ID is the identification number in the RIKEN database. The sublibrary ID is 000 for the standard library and 100, 104, or 304 for the subtracted/normalized cDNA libraries. In the subtracting or normalizing driver column, the value in brackets indicates the RoT used for each driver. The normalizing driver was always an aliquot of the starting RNA. The subtracting driver was mn1, minilibrary of liver, lung, brain, and placenta; mn2, minilibrary of liver, lung, brain, placenta; ms2, RIKEN nonredundant minilibrary (4000 clones); Nm2, RIKEN nonredundant minilibrary (1600 clones). The Method column indicates the method used to label the mRNA driver. Method 1, chemical end-biotinylation (Na10, biotin hydrazide long arm for cap and 3′-end of mRNA), enzymatic biotinylation (during RNA polymerase treatment using biotin-UTP), and biotin-psoralen (all RNAs); Method 2, label T biotin labeling kit only. Sequencing indicates the number of 3′-end sequencing passes. Species indicates the number of different species (clusters) obtained. Redundancy is the obtained redundancy (sequence/species). No. ESTs and No. NT indicate the number and percentage of sequences from the cDNA library that do not have relevant homology with EST or GenBank nucleotides (nonredundant GenBank +EMBL+DDBJ+PDB sequences), respectively. The percentage value refers to the number of clusters (species) and not to the number of total successful sequence. Coding indicates the percentage of clones among the mouse, complete mRNA that had the initiating ATG, as evaluated by 5′-end sequencing. Values in parentheses are when clones are <50 (but at least 10).

1994; Maruyama and Sugano 1994; Edery et al. 1995; Carninci at al. 1996, 1998; Carninci and Hayashizaki 1999), several of these methods involve PCR amplification. The associated preferential amplification of specific subpopulations of cDNAs adversely affects the discovery of rare and/or difficult-to-amplify cDNAs (Maruyama and Sugano 1994). The Cap-Trapper technique does not require PCR, thus leading to production of relatively deep and unbiased libraries. However, this method was not optimized for efficient discovery of rare, full-length cDNAs by one-pass sequencing without the use of tactics such as normalization and subtraction. It was generally considered problematic to prepare normalized/subtracted cDNA libraries that are at the same time full length (Rubin et al. 2000). Here we describe the strategy and methodology we developed to prepare normalized-subtracted cDNAs for genomic-scale, full-length cDNA discovery. Our technique greatly improves on the previous situation, in which normalized-subtracted cDNA libraries typically carried primarily incompletely synthesized cDNAs. This report shows for the first time the possibility of undertaking full-length, genomic-scale gene discovery by using a sequencing approach, as we show that the size of the libraries and the proportion of full-length cDNA inserts is very satisfactory.

Regarding the Size of Rare mRNAs

We repeatedly have observed that in alkali gels and checks of plasmid size, subtracted-normalized cDNAs seem to be longer than inserts from standard libraries. This finding is probably not an artifact of our normalization-subtraction method because we have efficiently subtracted long cDNAs by using magnetic beads in test experiments; the rate of new gene discovery confirmed this trend. In addition, the inserts of normalized-subtracted cDNA libraries are not shorter than those of standard libraries prepared with the same starting RNA. This result suggests that our method preserves the integrity of the cDNA after subtraction—as was confirmed by subsequent sequence analysis. This observation further suggests that the average length of the rarely expressed mRNAs is longer than the average length of the bulk cellular mRNA. Protocols that favor production and cloning of long, full-length cDNA inserts seem to increase the rate at which new genes are discovered in full-length cDNA libraries. Of particular concern in the generation of full-length cDNA libraries is the difficulty of constructing vectors that clone short and long cDNAs with the same efficiency and stability of a long plasmid vector during the propagation of the cDNA library. Notwithstanding these difficulties, we expect that the sequencing redundancy of a satisfactory full-length, normalized/subtracted cDNA library will be <1.5 when at least 7000 clones are sequenced and <2.0 for 15,000 clones.

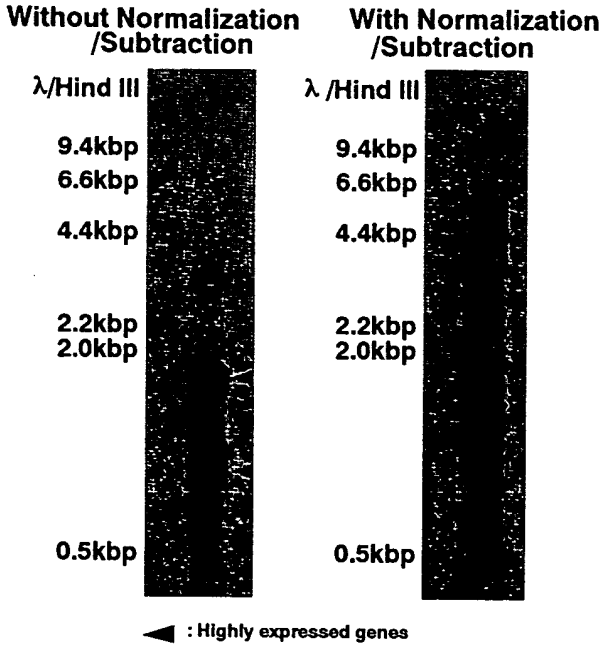


Figure 2 Visualization of removal of highly abundant full-length cDNAs. *Left,* second-strand cDNA prepared with control pancreas cDNA; *right,* cDNA prepared with an aliquot of the same pancreas cDNA after normalization/subtraction. Highly abundant cDNAs are indicated with an arrow and are removed in the normalized/subtracted cDNAs.

Relevance of Subtraction for Large-Scale Gene-Discovery Projects

Besides simple normalization, the key factor we are experiencing in the full-length cDNA gene discovery program is the importance of subtraction. Subtraction allows removal of already sequenced cDNAs as well as those that are predominantly expressed in other libraries. Subtraction helps keep the process of gene discovery efficient because resequencing of already-represented genes is reduced. Clearly this approach is facilitated by having the library production and sequencing centers in the same physical location so that feedback regarding clones to be used as drivers for subtraction occurs in a timely manner. Following this approach, we have prepared libraries subtracted with drivers corresponding to 30,000 different, previously sequenced cDNAs. In this situation, we are able to prepare libraries in which the rate at which new genes are discovered approached 25%–30% per successful sequencing reaction after clustering against a database of >80,000 3′-end sequences (not shown).

This rate of new gene discovery has to be considered extremely high after sequencing such a relevant part of 3′ ESTs. Subtraction removes 90%–95% of the mass of cDNA that might otherwise be represented in a library; this fact suggests that the rate of new gene discovery in an unsubtracted cDNA library would be 1.25%–3.0%. If we sequence 10,000 clones from a

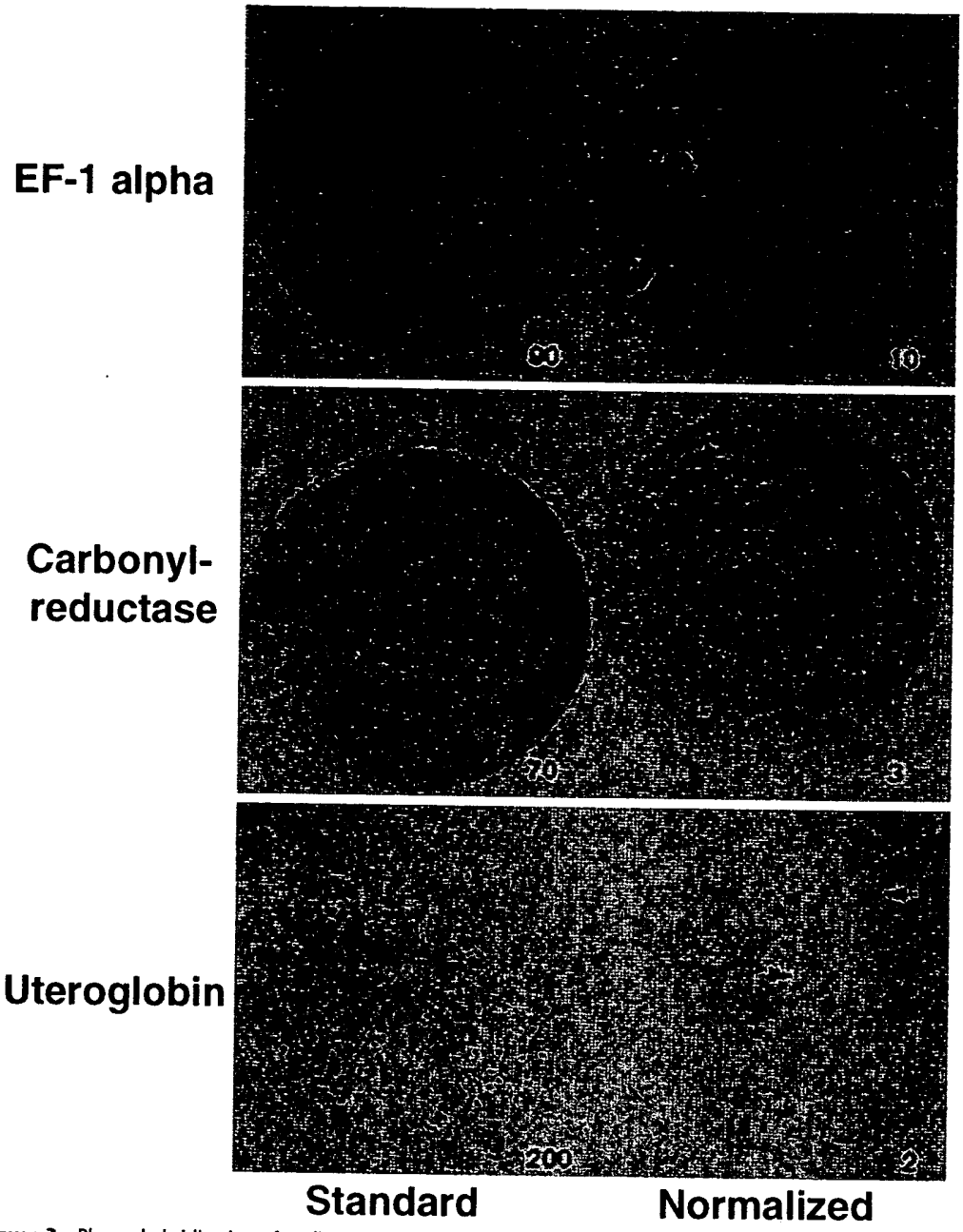


Figure 3 · Plaque hybridization of replicas containing control lung cDNA library (left) or a normalized lung cDNA library (right). In the right panel (normalized), an arrow indicates the plaque we have counted.

given efficiently subtracted cDNA library, we would be able to clone cDNAs that are expressed once in 100,000–200,000 events with a 50% probability. For cDNAs that are expressed even more infrequently, other strategies such as the preparation of cDNA libraries from defined subregions of tissues are required (P. Carninci, in prep.). We expect that our proposed methodology will be useful in the collection of the remaining human full-length cDNAs as well as the generation of cDNA encyclopedias for other organisms. Our method might further be used for applications that would benefit from normalized, full-length cDNA libraries, such as expression cloning.

METHODS

Harvest of mRNA and all other preparatory steps were completed as described previously (Carninci and Hayashizaki 1999).

cDNA Synthesis

In a total volume of 24 μL, we combined 5–10 μg mRNA, 5 μg of the first-strand primer containing the *Bam*HI and *Sst*I re-

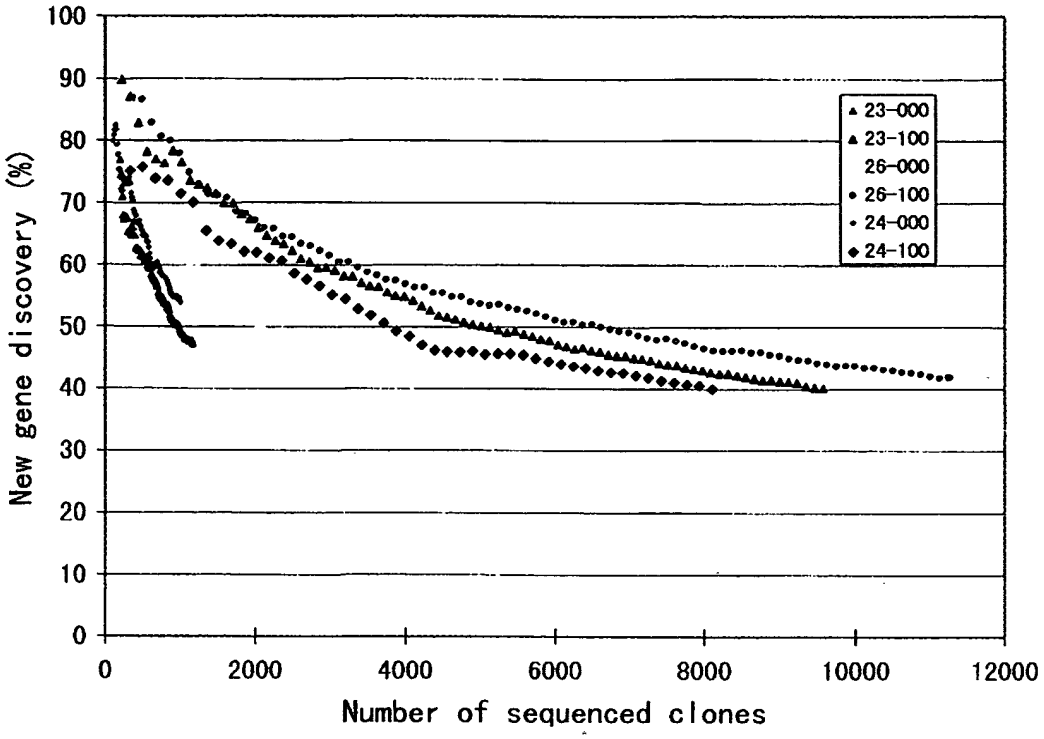


Figure 4 Sequencing redundancy (or the decrease in new gene discovery) increases sharply in standard cDNA libraries (-000 libraries), but in normalized/subtracted full-length cDNA libraries (-100 libraries), redundancy increases much more slowly. New genes (%) are referred as singleton (%) within a given cDNA library.

striction sites (5′-$(GA)_5$AAGGATCCAAGAGCTC$(T)_{16}$VN-3′), and 11.2 µL 80% glycerol. For liver and lung libraries and minilibraries, we instead used a primer containing an *XhoI* site (5′$(GA)_8$ACTCGAG$(T)_{16}$VN-3′), which generates inversely oriented cDNA libraries. The RNA-primer mixture was denatured at 65°C for 10 min. In parallel, we combined in a final volume of 76 µL; 18.2 µl 5× first-strand synthesis buffer; 9.1 µL 0.1 M DTT; 6.0 µL 10 mM (each) dTTP, dGTP, dATP, and 5-methyl-dCTP (instead of dCTP); 29.6 µL saturated trehalose (~80%, low metal content; Fluka Biochemika); and 10.0 µL Superscript II reverse transcriptase (200 U/µl). We placed 1.0 µl [α-^{32}P]dGTP in a third tube. The mRNA, glycerol, and primers were mixed on ice with the solution containing the Superscript, and an aliquot (20%) was quickly added to the tube containing the [α-^{32}P]dGTP. First-strand cDNA syntheses were performed in a thermocycler with a heated lid (e.g., MJ Research) according to the following program: step 1, 45°C for 2 min; step 2, gradient annealing: cool to 35°C over 1 min; step 3, complete annealing: 35°C for 2 min; step 4, 50°C for 5 min; step 5, 56°C for 60 min. Incorporation of radioactivity allowed us to estimate the yield of cDNA (Carninci and Hayashizaki 1999). The cDNA was treated with proteinase K, phenol/chloroform- and chloroform-extracted, and ethanol-precipitated by using ammonium acetate as the salt (Carninci and Hayashizaki 1999).

Table 2. Similar Presence of B1 Repeats in 200- and 300-Base Stretches in the 3-UTR of cDNAs from Subtracted/Normalized and Control Cap-Trapper cDNA Libraries

Library	Normalized subtracted sublibrary	Control sublibrary
Embryo 18	1.3% (143/10970)	1.6% (4/244)
Stomach	1.3% (114/8840)	0.6% (10/1606)
Tongue	1.2% (131/10974)	0.7% (10/1408)
ES cells	1.4% (209/15220)	0.9% (10/1089)
Embryo 13-liver	0.7% (37/5521)	1.0% (7/718)
Embryo 10	1.8% (225/12724)	1.3% (30/2239)
Embryo 12-head	1.8% (160/8873)	1.4% (7/490)

The number of clones positive for B1 sequences versus the total number of sequences analyzed is indicated in parentheses.

cDNA Biotinylation

Before biotinylation, the diol group of the cap and 3′-end of RNA were oxidized in a final volume of 50 µL, containing the resuspended first-strand cDNA, 66 mM sodium acetate (pH 4.5), and 5 mM $NaIO_4$. Samples were incubated on ice in the dark for 45 min. cDNA was then precipitated by adding 0.5 µL of 10% SDS, 11 µL NaCl, and 61 µl of isopropanol. After incubation in the dark on ice for 45 min or at −20° or −80°C for 30 min, the sample was centrifuged for 10 min at 15,000 rpm. Finally we rinsed the cDNA twice with 70% ethanol and resuspended it in 50 µL of water. Subsequently, the cap was biotinylated in a final volume of 210 µL by adding 5 µL M sodium acetate (pH 6.1), 5 µL 10% SDS, and 150 µL of 10 mM biotin hydrazide long-arm (Vector Biosystem).

After overnight (10–16 hr) incubation at room temperature (22°–26°C), the cDNA was precipitated by adding 75 µL 1

M sodium acetate (pH 6.1), 5 µL 5 M NaCl, and 750 µL absolute ethanol and incubated on ice for 1 hr or at −20° to −80°C for 30 min. The cDNA was pelleted by centrifugation at 15,000 rpm for 10 min; we then washed the pellet once with 70% ethanol and once with 80% ethanol. We resuspended the cDNA in 70 µL 0.1× TE (1 mM Tris [pH 7.5], 0.1 mM EDTA).

Capture and Release of Full-Length cDNA

We combined 500 µL of MPG-streptavidin beads and 100 µg DNA-free tRNA and incubated the mixture on ice for 30 min with occasional mixing. The beads were separated by using a magnetic stand for 3 min, and the supernatant was removed. The beads then were washed three times with 500 µL washing/binding solution (2 M NaCl, 50 mM EDTA [pH 8.0]).

At the same time, we added 1 U of RNase I (Promega) per microgram of starting mRNA to the cDNA sample in the buffer supplied by the manufacturer (final volume, 200 µL); the sample was incubated at 37°C for 15 min. To stop the reaction, we put the sample on ice and added 100 µg tRNA and 100 µL of 5 M NaCl. To capture the full-length cDNA, we combined the biotinylated, RNase I-treated cDNA and the washed beads, which were resuspended in 400 µL of the washing/binding solution. After mixing, the tube was gently rotated for 30 min at room temperature. Full-length cDNA remained on the beads, and the shortened cDNAs did not. The beads were separated from the supernatant on a magnetic stirrer. We gently washed the beads to remove the nonspecifically adsorbed cDNAs: Two washes with washing/binding solution; one with 0.4% SDS, 50 µg/mL tRNA; one with 10 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, 40 µg/mL tRNA, 10 mM NaCl, and 20% glycerol; and one with 50 µg/mL tRNA in water.

cDNA was released from the beads by adding 50 µL 50 mM NaOH, 5 mM EDTA and incubating for 10 min at room temperature with occasional mixing. The beads then were removed magnetically, and the eluted cDNA was transferred on ice to a tube containing 50 µL 1 M Tris-HCl, pH 7.5. The elution cycle was repeated once or twice with 50-µL aliquots of 50 mM NaOH, 5 mM EDTA until we recovered most of the cDNA (80%–90%, as measured by monitoring the radioactivity with a handheld monitor) from the beads.

To remove traces of RNA that could later interfere with the biotinylated RNA driver, we quickly added 100 µL of 1 M Tris-HCl, pH 7.0, and 1 µL RNase I (10U/µL) to the recovered cDNA on ice; the sample then was incubated at 37°C for 10 min. The cDNA was treated with proteinase K, phenol/chloroform-extracted, and back-extracted. We then added 2–3 µg glycogen and ethanol-precipitated the sample in a siliconized tube. Alternatively, the sample can be concentrated by using one round of ultrafiltration with a Microcon 100 (Millipore) for 40–60 min at 2000 rpm. If ethanol precipitated, the cDNA could be redissolved in 20 µl of 0.1 × TE.

CL-4B Spin-Column Fractionation of cDNA

We treated the cDNA samples with CL-4B chromatography (Carninci and Hayashizaki 1999) or an S-400 spin column (Amersham-Pharmacia) essentially as described by the manufacturer.

Oligo-dG Tailing of the First-Strand cDNA

We combined the cDNA sample, 5 µL of the 10× TdT buffer (2 M potassium cacodylate [pH 7.2], 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol), 5 µL of 50 µ M dGTP, 5 µL of 10 mM $CoCl_2$, and 40 U terminal deoxynucleotidyl transferase in a final volume of 50 µL. Samples were incubated at 37°C for 30 min. At the end, reaction was stopped with EDTA 20 mM, cDNA digested with proteinase K, extracted with phenol chloroform, and ethanol precipitated. Sample was finally redissolved in TE. We checked the tail length as described (Carninci et al. 1999), after which the cDNA was used in a second-strand synthesis for use in check libraries (see later) or underwent normalization/subtraction.

Normalization Drivers

mRNA drivers comprising an aliquot of the starting mRNA were called "normalizing drivers." To calculate the concentration of the normalizing driver, we approximated the ribosomal/structural RNA contamination in the starting mRNA by assuming that the incorporation rate of the first-strand synthesis reflected the actual mRNA concentration, thus assuming 100% efficiency of priming and elongation. Assuming that the proportion of mRNA converted to first-strand cDNA corresponded to the effective mRNA concentration, we omitted accounting for less than full-length cDNAs—usually not all of the mRNA is primed. A slight excess of normalization driver was unlikely to interfere with the normalization process as dramatically as would a paucity of driver. Therefore, we assumed that the amount of mRNA in the sample was the same as the quantity of first-strand cDNA produced.

Subtraction Drivers

Subtracting drivers comprised bulk run-off transcripts prepared from cloned minilibraries and rearrayed libraries from the nonredundant RIKEN cDNA encyclopedia by using T7 and T3 RNA polymerases.

Minilibraries contain ~1000–2000 clones of cDNA deriving from a previous normalization experiment. By adapting the standard protocol, we prepared minilibraries from the captured aliquot (abundant cDNA fraction) that was the byproduct of normalization experiments. After normalization, the abundant cDNA fraction was removed from the beads with 50 mM NaOH/5 mM EDTA; after neutralization, second-strand cDNA was prepared. Cloning was accomplished in a way analogous to that previously described (Carninci and Hayashizaki 1999). Plasmid was then bulk-excised, and 1000–2000 clones per minilibrary were amplified on agarose/ampicillin. For driver preparation, we plated 20,000–50,000 colonies on SOB-agarose/ampicillin and incubated the plates overnight at 37°C. We scraped bacterial cells from the plate in the presence of resuspension solution (Wizard DNA extraction kit; Promega,) and later followed the manufacturer's protocol.

Preparation of the Nonredundant cDNA Library Driver

Single clones from the full-length cDNA encyclopedia (http://genome.rtc.riken.go.jp/) were rearrayed for the subtraction. From 384-well plates, rearrayed cDNAs were then plated on SOB-agarose/ampicillin plates. Plasmid extraction, DNA cleavage, and RNA preparation was performed as for minilibraries.

We digested the extracted plasmid at the 3′-end of the multiple cloning site by using *PvuI* when the minilibrary was cloned with *XhoI* at the 3′-end site or *SstI* when the library was cloned in the *SstI* site. RNA was synthesized by using either T3

or T7 RNA polymerase (Life Technologies), depending on the map of the construct used to prepare the driver, to prepare sense run-off RNAs. We used T3 polymerase for *Pvu*I-cleaved minilibraries (up to 14), and T7 polymerase for *Sst*I-cleaved minilibraries (15 and following). RNA was prepared by using RNA polymerases (Life Technologies) according to manufacturer's instructions. Extensive digestion with 1–2 µL DNaseI (RQ1, RNase-free, Promega) was performed for 30 min. Proteinase K digestion was then performed, followed by extraction with phenol/chloroform and chloroform, and the cDNA was precipitated.

Biotin Labeling of Normalizing/Subtracting RNA Drivers

To further clean up RNA drivers before labeling, we used the RNeasy kit (QIAGEN) according to the instructions of the manufacturer. Subsequently, we used the Mirus nucleic acid biotinylation kit (Panvera) essentially as described by the manufacturer. For instance, 10 µg of the RNA mix was labeled by combining it with 10 µL of Label IT reagent and 10 µL of labeling buffer A, in a final volume of 100 µl. We incubated the reaction at 37°C for 1 hr, after which we precipitated the biotinylated RNA by adding 1/20-volume of 5M NaCl and two volumes of 99% ethanol. After standard ethanol precipitation, the pellet was washed once with 80% ethanol, resuspended in 20 µL of 1× Mirus labeling buffer A, and stored at −80°C until used. Alternatively, mRNA was labeled by using the psoralen-biotinylation kit (Ambion) according to the instructions of the manufacturer.

Normalization/Subtraction

The RNA drivers and cDNA were deproteinated by using proteinase K followed by phenol/chloroform extraction, chloroform extraction, and ethanol precipitation. Oligo-dG-tailed cDNA was used as a substrate, which was mixed with the RNA drivers and blocking oligonucleotides (biotin-dG_{16}) to hybridize to the C-stretch present in the subtracting driver and with oligo-dT primer to block the polyA sequences. Hybridization was carried out at RoT values of 5–500, depending on the experiment, in a buffer containing 80% formamide (from a deionized stock), 250 mM NaCl, 25 mM HEPES (pH 7.5), and 5 mM EDTA. Hybridization was carried out at 42°C in a dry oven; even volumes as small as 5 µL did not require mineral-oil overlays. After hybridization, we precipitated the sample by adding 2.5 volumes of absolute ethanol and incubated it for 30 min on ice. The sample was centrifuged for 10 min at 15,000 rpm and washed once with 70% ethanol; we carefully resuspended the cDNA in 10 µL of water on ice.

Removal of the Hybrid

In parallel, we prepared 50 µl CPG magnetic beads for each 1 µg of biotinylated driver RNA; 5 µL beads could bind >400 ng of biotinylated driver. To each 50 µL of beads, we added 10 µg tRNA as a blocking agent, then incubated the beads at room temperature for 10–20 min or on ice for 30–60 min with occasional shaking. We used a magnetic stand to remove the beads, which we washed three times with a large excess of 1 M NaCl, 10 mM EDTA and resuspended them in a volume of 1 M NaCl, 10 mM EDTA equivalent to the original volume of the bead suspension.

We mixed the blocked beads with the redissolved tester/driver mixture and incubated the entire sample at room temperature for 15 min with occasional gentle mixing. After removing the beads by using a magnetic stand for 3 min, we recovered the supernatant, which contained the single-strand normalized/subtracted cDNA. The beads were washed once with excess volume of binding buffer (1 M NaCl, 10 mM EDTA) to recover any remaining ssDNA. We measured the radioactivity of the labeled samples before and after the procedure in order to estimate the yield of normalization/subtraction.

To concentrate the cDNA solution to ~50 µL, we used Microcon 100 ultrafiltration as described by the manufacturer (Millipore). Subsequently, the cDNA was pelleted by using the standard isopropanol procedure; the pellet was resuspended in 44 µL of 0.1 × TE, to which 5 µL of RNase I buffer and 1 U RNase I were added, in a volume of 50 µL. Samples were then incubated for 20 min at 37°C, after which we added 400 µL of 0.2% SDS to inactivate the RNase I. Traces of degraded RNAs, blocking oligonucleotide, SDS, and buffer were removed by ultrafiltration with a Microcon 100 filter at 2000 rpm and 25°C until the volumes were reduced to <20 µL. The samples were desalted by adding 400 µl of 0.1 × TE then centrifuging as above for a total of three washes. We recovered the cDNA by inverting the filter in a new tube and centrifuging at 9000 rpm for 1 min.

Second-Strand cDNA Synthesis

For normalized/subtracted cDNA, the standard control libraries, and the minilibraries, the second-strand synthesis and cloning steps were the same. The *Xho*I-containing primer 5′-$(GA)_7$TTCTCGAGTTAATTAAATTAATC$_{13}$-3′ was prepared and purified by using standard techniques, as was the first-strand cDNA primer. For the lung and liver libraries and minilibraries, the *Sst*I-containing primer 5′-$(GA)_9$GAGCTCACT-AGTTTAATTAAATTAATC$_{11}$-3′ was used as the second-strand primer. To prepare the second-strand reaction, we mixed the oligo-dG-tailed cDNA with 6 µl of 100 ng/µL second-strand primer adapter, 6 µL of EX-Taq second-strand buffer (Takara), and 6 µL 2.5 mM (each) dNTPs. Hot-start priming then was performed by adding 3 µl of 5 U/µL ExTaq polymerase (Takara) at 65°C in a thermocycler. After mixing, the annealing temperature was reached by a negative ramp to 45°C for the *Xho*I primer and to 35°C for the *Sst*I primer. After 10 min at the annealing temperature, the second-strand cDNA was extended during incubation at 68°C for 20 min. The annealing-extension cycle was repeated once more, followed by a final elongation step at 72°C for 10 min. At the beginning of the hot-start, we mixed a 5 µL aliquot with 0.5 µL of [α^{32}P]dGTP or [α^{32}P]dCTP to follow the incorporation. We used the labeled aliquot at the end of reaction to visualize the cDNA and to calculate the second-strand yield (Carninci and Hayashizaki 1999).

cDNA Cloning

Second-strand cDNA was treated with proteinase K, extracted with phenol-chloroform and chloroform, and ethanol-precipitated according to standard procedures. We then cleaved the cDNA by using 25 U/µg each of *Sst*I and *Xho*I (lung and liver libraries and Lib.18–31) or *Bam*HI and *Xho*I (Lib.32–64). After the digestion, cDNA was treated with proteinase K, extracted with phenol-chloroform, and purified over a CL-4B spin column (Pharmacia). After ethanol precipitation, we cloned the cDNA essentially as described (Carninci and Hayashizaki 1999). The vector for cloning the cDNA Lib.32–64 will be described elsewhere (P. Carninci, in prep.).

Other Methods

Plaque hybridization was performed by using a random primer according to standard protocols (Sambrook et al. 1989). Alkali electrophoresis was performed as described (Sambrook et al. 1989). All autoradiography signals were visualized by using the Bas 2000 imaging system (Fuji).

Bacteria were picked with commercially available picking machines (Q-bot and Q-pix; Genetics, UK) and transferred to 384-microwell plates. Duplicate plates were used to prepare plasmid DNA. For plasmid DNA, 384-well plate were divided and grown in 4 × 96 deep well plates. After overnight growth, plasmids were extracted either manually (Itoh et al. 1997) or automatically (Itoh et al. 1999). Sequences typically were run on the RISA sequencing instrument (K. Shibata, in prep.); a few sequences were generated by using the Perkin Elmer-Applied Biosystems ABI 377. Sequencing primers were the M13 forward and reverse primers, and the main sequencing operation will be described in detail elsewhere (P. Carninci, in prep.).

Sequences for clustering were analyzed as follows. The poly-T (for 3′-end) and C-stretch (for 5′-end) regions were trimmed from the one-pass sequences. From the trimmed sequences, we selected 100-bp sequences to use as tag sequences. We used BLAST 2.0.9 to search for homology between the new tag and the database of nonredundant 100-bp tag sequences; sequences having BLAST parameters of $E = 1.0e^{-25}$ or lower were clustered together. When the tag sequence was not in the database, the tag was added to the database. When the database contained the tag, it was added to the member of identical group of the TAG. If the tag was found in the database and at the same time the shift was <10 bases, the overlap was >80 bases, with >90% identity in the overlap, the sequences were grouped together. In addition, the algorithm categorized sequences within the library as "new," "nonredundant," or "redundant," according to the previously defined criteria.

ACKNOWLEDGMENTS

We dedicate this work to Yuichi Sugahar, coauthor of this article, who died prematurely in an accident. We thank Claudio Schneider for multiple discussions and encouragement; Tomoko Hirozane, Toshiyuki Shiraki, and Kenjiro Sato for excellent technical contributions; and all members of the Genome Science Laboratory for collecting the data. This study has been supported by Special Coordination Funds and a Research Grant for the RIKEN Genome Exploration Research Project, CREST (Core Research for Evolutional Science and Technology), and ACT-JST (Research and Development for Applying Advanced Computational Science and Technology) of Japan Science and Technology Corporation (JST). Y.H. was funded by the Science Technology Agency in Japanese Government. This work was also supported by a Grant-in-Aid for Scientific Research on Priority Areas and Human Genome Program, from the Ministry of Education, Science and Culture, and by a Grant-in-Aid for a second Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare to Y.H.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Adams, M., Kelley, J., Gocayne, J., Dubnick, M., Polymeropoulos, M., Xiao, H., Merril, C., Wu, A., Olde, B., Moreno, R., et al. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* **252:** 1651–1656.

Adams, M., Kerlavage, A., Fleischmann, R., Fuldner, R., Bult, C., Lee, N., Kirkness, E., Weinstock, K., Gocayne, J., White, O., et al. 1995. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides cDNA sequence. *Nature* **377:**(Suppl.) 3–174.

Anderson, M., and Young, B.D. 1985. Quantitative analysis of solution hybridisation. In *Nucleic Acids Hybridisation, a practical approach*, pp. 73–111. IRL Press, Oxford.

Barr, F. and Emanuel, B. 1990. Application of a subtraction hybridization technique involving photoactivatable biotin and organic extraction to solution hybridizationanalysis of genomic DNA. *Anal. Biochem.* **186:** 369–373.

Bonaldo, M., Lennon, G., and Soares, M. 1996. Normalization and subtraction: Two approaches to facilitate gene discovery. *Genome Res.* **6:** 791–806.

Carninci, P. and Hayashizaki, Y. 1999. High-efficiency full-length cDNA cloning. *Methods Enzymol.* **303:** 19–44.

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. 1998. Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. *Proc. Natl. Acad. Sci.* **95:** 520–524.

Carninci, P., Kvam, C., Kitamura, A., Ohsumi, T., Okazaki, Y., Itoh, M., Kamiya, M., Shibata, K., Sasaki, N., Izawa, M., et al. 1996. High-efficiency full-length cDNA cloning by biotinylated CAP trapper. *Genomics* **37:** 327–336.

Diatchenko, L., Lau, Y., Campbell, A., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E., et al. 1996. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci.* **93:** 6025–6030.

Edery, I., Chu, L., Sonenberg, N., and Pelletier, J. 1995. An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). *Mol. Cell Biol.* **15:** 3363–3371.

Fargnoli, J., Holbrook, N., Fornace Jr., A. 1990. Low-ratio hybridization subtraction. *Anal. Biochem.* **187:** 364–73.

Hillier, L., Lennon, G., Becker, M., Bonaldo, M., Chiapelli, B., Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., et al. 1996. Generation and analysis of 280,000 human expressed sequence tags. *Genome Res.* **6:** 807–828.

Hirotsune, S., Takahara, T., Sasaki, N., Hirose, K., Yoshiki, A., Ohashi T., Kusakabe, M., Murakami, Y., Muramatsu, M., Watanabe, S., et al. 1995. The reeler gene encodes a protein with an EGF-like motif expressed by pioneer neurons. *Nat. Genet.* **10:** 77–83.

Kato, S., Sekine, S., Oh, S., Kim, N., Umezawa, Y., Abe, N., Yokoyama-Kobayashi, M., and Aoki, T. 1994. Construction of a human full-length cDNA bank. *Gene* **150:** 243–250.

Itoh, M., Carninci, P., Nagaoka, S., Sasaki, N., Okazaki, Y., Ohsumi, T., Muramatsu, M., Hayashizaki, Y. 1997. Simple and rapid preparation of plasmid template by a filtration method using microtiter filter plates. *Nucleic Acids Res.* **25:** 1315–1316.

Itoh, M., Kitsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y. et al. 1999. Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* **9:** 463–470

Marra, M., Hillier, L., Kucaba, T., Allen, M., Barstead, R., Beck, C., Blistain, A., Bonaldo, M., Bowers, Y., Bowles, L., et al. 1999. An encyclopedia of mouse genes. *Nature Genet.* **21:** 191–194.

Maruyama, K. and Sugano, S. 1994. Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138:** 171–174.

Rubin, G., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M., and Harvey, D. 2000. A *Drosophila* complementary cDNA resource. *Science* **287:** 2222–2224.

Sambrook, J., Fritsch, E., and Maniatis, T. 1989. *Molecular Cloning.*

Cold Spring Harbor laboratory press, Cold Spring Harbor, NY.
Sasaki, Y., Ayusawa, D., and Oishi, M. 1994. Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system. *Nucleic Acids Res.* **22:** 987–992.
Sasaki, N., Nagaoka, S., Itoh, M, Izawa, M., Konno, H., Carninci, P., Yoshiki, A., Kusakabe, M., Moriuchi, T., Muramatsu, M., et al. 1998. Characterization of gene expression in mouse blastocyst using single-pass sequencing of 3995 clones. *Genomics* **49:** 167–79.
Soares, M., Bonaldo, M., Jelene, P., Su, L., Lawton, L., and Efstratiadis, A. 1994. Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci.* **91:** 9228–9232.
Takahashi, N. and Ko, M. 1994. Toward a whole cDNA catalog: Construction of an equalized cDNA library from mouse embryos. *Genomics* **23:** 202–210.
Tanaka, T., Ogiwara, A., Uchiyama, I., Takagi, T., Yazaki, Y., and Nakamura, Y. 1996. Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing. *Genomics* **35:** 231–235.

Received April 20, 2000; accepted in revised form July 24, 2000.

RE1.IGS
mail/mm
MPL

STIC-ILL

Fr m: Lacourciere, Karen
Sent: Wednesday, July 24, 2002 10:31 AM
To: STIC-ILL
Subject: ill order

Please provide the following reference for use in examining 09/836,410

FASEB Journal, (March 12, 1999) Vol. 13, No. 4 PART 1, pp. A531.

IOVS, (March 15, 2001) Vol. 42, No. 4, pp. S94. print.

INVESTIGATIVE OPHTHALMOLOGY AND VISUAL SCIENCE, (2001 Nov) 42 (12) 3000-7.

Meth. Enzymol. 303, 19-44, (1999).

Genome Res. 10 (10) 1617-1630 (2000)

Genome res. 10 (11) 1757-1771 (2000)

Thank-You!

Karen A. Lacourciere Ph.D.
CM1 11D09 GAU 1635
(703) 308-7523

Suppressed Expression of Tubedown-1 in Retinal Neovascularization of Proliferative Diabetic Retinopathy

Robert L. Gendron,[1] *William V. Good,*[2] *Lisa C. Adams,*[1] *and Hélène Paradis*[1]

Purpose. Retinal neovascularization occurring as a complication of diabetes mellitus can cause vision loss and blindness. The identification and study of novel genes involved in retinal angiogenesis may define new targets to suppress retinal neovascularization in diabetes and other ocular diseases. A novel acetyltransferase subunit, tubedown-1 (tbdn-1), has been isolated, the expression of which is regulated during blood vessel development. Tbdn-1 is not detected in most adult vascular beds but persists at high levels in the adult ocular vasculature. The purpose of this study was to gain insight into the possible role of tbdn-1 in retinal blood vessels by characterizing its expression patterns in adult homeostasis and in retinal neovascularization associated with diabetes.

Methods. Western blot analysis and immunohistochemistry were performed to study the expression patterns of tbdn-1 during adult homeostasis in normal human retinas, in a model of choroid-retina endothelial capillary outgrowth in vitro, and in retinas showing neovascularization in patients with proliferative diabetic retinopathy (PDR).

Results. In adults during homeostasis, tbdn-1 was expressed highly in normal endothelium of retinal and limbic blood vessels. Tbdn-1 was also expressed in RF/6A, a rhesus macaque choroid-retina-derived endothelial cell line. In an in vitro model system using the RF/6A cell line, tbdn-1 expression was downregulated during the outgrowth of these cells into capillary-like structures on a reconstituted basement membrane matrix. Similar to this in vitro model, tbdn-1 expression is specifically suppressed in the endothelial cells of blood vessels and capillary fronds in vivo in both the neural retinal tissue and in preretinal membranes in eyes of patients with PDR.

Conclusions. High levels of expression of tbdn-1 are associated with ocular endothelial homeostasis in adults. Conversely, low levels of tbdn-1 expression are associated with endothelial capillary outgrowth in vitro and retinal neovascularization in vivo. Because the tbdn-1 acetyltransferase subunit is a member of a family of regulatory enzymes that are known to control a range of processes, including cell growth and differentiation, through posttranslational modification, the current results support a hypothesis that tbdn-1 may be involved in maintaining homeostasis and preventing retinal neovascularization. (*Invest Ophthalmol Vis Sci.* 2001;42:3000-3007)

Diseases involving ocular neovascularization can cause visual loss and blindness. Ocular disease associated with diabetes mellitus is initially characterized by retinal ischemia, which progresses to a proliferative stage involving both neovascularization of the retina, optic disc, or iris and fibrosis.[1,2] In proliferative diabetic retinopathy (PDR), high levels of proangiogenic factors are thought to lead to increased neovascularization, which contributes to a positive feedback cycle of fibrovascular growth, retinal dysplasia, scarring, and eventual retinal detachment.[3] A range of angiogenic growth factors (VEGF, bFGF, and insulin-like growth factor [IGF]-1), integrins and extracellular matrix components probably contribute to and have been associated with pathologic neovascularization in PDR.[3-5] However, increased production of VEGF in the retina may be a determining factor in the later proliferative neovascularization that leads to pathologic sequelae in later stages of PDR.[5-7] The current effective treatment for PDR involves ablative therapy that can cause complications (retinal vein occlusion, loss of visual acuity, vitreous hemorrhage) or even sometimes fails altogether.[1,2]

Treatments specifically targeting either VEGF and its receptors or specific integrins have been found effective in reducing but not abolishing retinal neovascularization in animal models.[8-12] Because a range of angiogenic factors is probably involved in the microenvironment promoting retinal blood vessel proliferation, the targeting of a single factor for antiangiogenic therapy may not completely counter the neovascularization in PDR. The characterization of common regulators that act downstream of these angiogenic signals mediating retinal neovascularization are key to identifying targets that could have a more global effect on controlling retinal neovascularization.

At present, there is little known about the intracellular regulatory pathways controlling retinal neovascularization and the presumed disturbances in such pathways during PDR. Elucidation of these regulatory pathways and the identification of the associated molecular effectors could reveal potential targets for blocking neovascularization and restoring normal function to the diseased retina. We have isolated a novel gene that we named tubedown-1 (*tbdn-1*), because it is downregulated during the formation of capillary structures in IEM vascular endothelial cells in vitro and during the development of most vascular beds in vivo.[13] *Tbdn-1* encodes a novel 69-kDa polypeptide associated with an acetyltransferase activity.[13] *Tbdn-1* displays homology with the previously characterized yeast N-terminal acetyltransferase subunit NAT1 and contains other motifs suggesting a regulatory function.[13] In contrast to most vascular structures in adults, *tbdn-1* expression persists in ocular vascular endothelium in adulthood. In the present study, we investigated the expression regulation of *tbdn-1* in normal and diseased eyes to better understand the potential role of this novel regulatory protein in retinal vessel homeostasis.

From the [1]Department of Pediatrics, University of Cincinnati College of Medicine, Division of Hematology and Oncology, The Children's Hospital Research Foundation, Cincinnati, Ohio; and the [2]Smith-Kettlewell Eye Research Institute, San Francisco, California.

Supported by Grants EY12827-01 (RLG, HP) and EY00384-03 and EY12472-02 (WVG) from the National Institutes of Health; and a Chair's Development Fund grant from the Childrens Oncology Group (RLG, HP).

Submitted for publication January 17, 2001; revised July 17, 2001; accepted August 6, 2001.

Commercial relationships policy: N.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Robert L. Gendron, Department of Pediatrics, University of Cincinnati College of Medicine, Division of Hematology and Oncology, The Children's Hospital Research Foundation, 3333 Burnet Avenue, Cincinnati, OH 45229. rlgendron@chmcc.org

Investigative Ophthalmology & Visual Science, November 2001, Vol. 42, No. 12

Materials and Methods

Cell Culture

RF/6A rhesus macaque choroid-retina endothelial cells[14,15] were obtained from the American Type Culture Collection (Manassas, VA) and were grown in DMEM supplemented with 10% fetal bovine serum (FBS) plus 2 mM glutamine and nonessential amino acids. The spontaneously immortalized RF/6A choroid-retina endothelial cell line, derived from the choroid-retina of a rhesus macaque, retains the expression of endothelial markers[14,15] including the VEGF receptor type-2 tyrosine kinase (VEGFR-2; our unpublished observation, 2000). Human umbilical vein endothelial cells (HUVECs) were obtained from Clonetics (San Diego, CA) and grown in DMEM plus 10% FBS, 2 mM glutamine, 1 ng/ml bFGF, and a mixture of insulin, transferrin, and selenium (Gibco, Rockville, MD). The IEM cell line, from which *tbdn-1* was originally isolated, was grown as previously described[16] and was initially derived by immortalizing differentiation products of embryonic stem cell cultures using simian virus (SV)40 large T antigen. The IEM line expresses endothelial markers and can be induced to form capillary structures in a synthetic basement matrix (Matrigel; Collaborative Research, Inc., Bedford, MA) after induction with bFGF and leukemia inhibitory factor.[16] IEM cells can also contribute to vascular structures in embryonic chimeras in vivo after blastocyst injections.[16] *Tbdn-1* RNA and protein become downregulated as IEM cells differentiate into capillaries on synthesized basement membrane.[13]

Capillary Formation

For capillary induction, cultures of RF/6A cells were treated with 10 ng/ml bFGF plus 10 ng/ml VEGF for 48 hours before being transferred to the synthetic basement membrane for a further 96 hours for capillary formation, as previously described.[13,16,17] RF/6A capillary colonies were collected by gently lifting the colonies, together with the membrane on which they were growing, from the culture dishes with a fine spatula. Control cultured RF/6A cells were harvested from tissue culture dishes by scraping the cells from the dishes and were collected by gentle centrifugation. The pellets of cultured RF/6A cells and the RF/6A capillary colonies were then fixed in 4% buffered paraformaldehyde and immobilized by embedding in small blocks of low-melting-temperature agarose. The agarose blocks containing the pellets of cultured RF/6A cells and the RF/6A capillary colonies were next fixed in 4% paraformaldehyde and embedded in paraffin blocks for histologic processing and analysis.

Anti-tbdn-1 Antibody

As described in our initial report of the cloning and characterization of *tbdn-1*,[13] an anti-tbdn-1 IgY antibody (Ab1272) was generated by immunizing chickens with a keyhole limpet hemocyanin (KLH)-conjugated 10mer peptide sequence in the *tbdn-1* open reading frame. The peptide sequence used was MDEAQALDTA (tbdn-1 amino acids 160-170). IgY was isolated to 90% purity from preimmune and immune egg yolks using an extraction agent (Eggstract; Promega, Madison, WI). We have previously demonstrated the specificity of Ab1272 for detecting tbdn-1 protein in IEM cell lysates by Western blot analysis and in tissue sections by immunohistochemistry.[13]

Tissue Specimens and Immunohistochemistry

Immunocytochemistry was performed on paraformaldehyde-fixed, paraffin-embedded sections of cultures of untreated RF/6A cells, RF/6A capillary colonies, and human eye tissues to detect tbdn-1 and endothelial and pericyte marker expression. Four normal human adult eye specimens and five specimens from patients with PDR were studied. All human eye tissue specimens were obtained under the approval of the Institutional Review Boards of the Smith Kettlewell Eye Research Institute, San Francisco, and Childrens Hospital Medical Center, Cincinnati. All research on human specimens followed the tenets of the Declaration of Helsinki at all times. Specimens of human eyes were obtained either from the University of San Francisco Department of Ophthalmologic Pathology or procured for us from certified eye banks through services of the National Disease Research Interchange (Philadelphia, PA). Normal control eye specimens were obtained as either whole globes or posterior poles from donors with no history of eye disease. Procurement criteria for PDR specimens received through the National Disease Research Interchange were that donors must have diabetic retinopathy, must be 21 to 100 years of age, but could be of either sex and any race.

Eyes of patients with diabetes (either whole globes or posterior poles) were procured at autopsy within 8 hours of death and formalin fixed within 12 hours of death. Formalin fixed eyes were processed and embedded in paraffin blocks. The health history and details of the diabetic retinopathy eye specimens analyzed were as follows: The first specimen was from an enucleation in a 74-year-old man with long-standing insulin-dependent diabetes mellitus (IDDM) with a history of unresolved PDR. The second specimen was from a 78-year-old man with long-standing IDDM and a history of unresolved diabetic retinopathy who had died of myocardial infarction. The third specimen was from a 70-year-old woman with long-standing IDDM with a history of unresolved diabetic retinopathy who had died of acute myocardial infarction. The fourth specimen was from a 60-year-old man with long-standing IDDM and a history of unresolved diabetic retinopathy who had been found unresponsive, had undergone attempted cardiopulmonary resuscitation, and was declared dead on arrival at the hospital. The fifth specimen was from a 62-year-old man with long-standing IDDM and a history of unresolved diabetic retinopathy who died of acute renal failure.

Specimens were embedded in paraffin blocks and were sectioned. They were then deparaffinized, rehydrated, and subjected to immunohistochemistry. All conditions and procedures for processing RF/6A cells and RF/6A capillary colonies were identical. After a 1-hour blocking step in 2% normal goat serum, sections were incubated with either a 1:100 dilution of chicken anti-tbdn-1 IgY (Ab1272[13]) or an equal concentration of preimmune IgY. For an endothelial cell marker, rabbit anti-von Willebrand factor (Dako, Glostrup, Denmark) was used for labeling endothelial cells in blood vessels in adjacent sections. An anti-α-smooth muscle actin (ASMA) monoclonal antibody directly conjugated to alkaline phosphatase (Sigma, St. Louis, MO) was used in conjunction with the anti-tbdn-1 IgY antibody to double label sections of normal human eyes for simultaneous localization of tbdn-1 and pericytes. Anti-tubulin mouse monoclonal antibody (Sigma) was used as a positive ubiquitous staining control for RF/6A cells and capillary colonies. After a rinse in phosphate-buffered saline (PBS), reactions were developed using the appropriate alkaline phosphatase–conjugated, species-specific secondary reagents (anti-rabbit IgG, anti-mouse IgG, or anti-chicken IgY; Promega). Red color reactions were generated using naphthol-AS-MX phosphate in the presence of fast red and levamisole (to block endogenous tissue alkaline phosphatase activity). In double-labeling experiments, the anti-tbdn-1 reaction was developed first using anti-IgY horseradish peroxidase and a diaminobenzidine (DAB) substrate kit (Sigma) to yield a dark brown color reaction for tbdn-1 expression, whereas the alkaline phosphatase anti-ASMA reaction was developed immediately after using fast red and levamisole, as stated above. Anti-tbdn-1/anti-ASMA double-labeled reactions were slightly overdeveloped to enable clearly revealing the locations of both epitopes. Slides were counterstained lightly using a 0.5% aqueous solution of methyl green. Sections were then rinsed, dried, and mounted (Permount, Fisher, Pittsburgh, PA) before viewing and photography using a microscope-mounted digital camera (DC120; Eastman Kodak, Rochester, NY). Differences in immunohistochemical staining of tbdn-1 were quantitatively analyzed by measuring the total area of red chromogen in high-power fields of identical dimensions sampled from the retinal areas in normal and PDR specimens. Measurements were made using the magic wand and histogram command tools of an image-management program (Photoshop; Adobe, Mountain View, CA) run on a computer (Macintosh G3; Apple Computer, Cupertino, CA),

as described in a previously published method.[18] Results are expressed as mean red chromogen pixels per high-power field ± SEM.

Western Blot Analysis

Cell lysates were prepared using Triton X-100 lysis buffer (50 mM Tris [pH 8.0], 150 mM NaCl, 1% Triton X-100) supplemented with protease inhibitors (1 mM phenylmethylsulfonyl fluoride [PMSF], 0.3 U/ml aprotinin, and 10 μg/ml leupeptin) and phosphatase inhibitors (1 mM sodium orthovanadate, 25 mM sodium fluoride, and 10 mM β-glycerophosphate). Lysates were clarified by centrifugation, the protein concentration was quantified and samples analyzed by SDS-PAGE. Western blot analysis was performed by standard procedures using chemiluminescence detection (ECL Plus reagent; Amersham Pharmacia Biotech, Piscataway, NJ), except for low-salt buffer (25 mM NaCl) conditions for Ab1272 incubations and washes. For experiments demonstrating the specificity of the Ab1272 antibody in Western blot analysis of RF/6A cells, RF/6A cell clones overexpressing *tbdn-1* cDNA sequences 1-1413 in an antisense orientation were generated using zeocin selection from the pcDNA3.1/Zeo vector (Invitrogen, San Diego, CA). Lysates were prepared from several of these antisense tbdn-1 RF/6A transfectants and from parental RF/6A cells, as described earlier, and then used in Western blot analysis experiments for testing the specificity of Ab1272 in detecting tbdn-1 in RF/6A cells.

Results

Tbdn-1 Expression in Endothelial Cells

We first made a comparison of tbdn-1 expression in endothelial cell lines from different species. To establish that tbdn-1 is equally detectable by Ab1272 anti-tbdn-1 antibody in primate and human retinal endothelial cells as it is in mouse vascular endothelial cells,[13] we first performed Western blot analysis using the Ab1272 antibody on whole-cell lysates prepared from the rhesus macaque RF/6A choroid-retina endothelial cell line,[14,15] the mouse IEM embryonic endothelial cell line,[13,16] and HUVECs. Western blot analysis indicated the presence of a 6-kDa tbdn-1 protein band in all these endothelial cell lines (Fig. 1, left). As we have described previously, IEM cells display a 69-kDa doublet that could correspond to acetylated and nonacetylated forms of tbdn-1.[13] Furthermore, Ab1272 Western blot analysis of several RF/6A cell clones stably overexpressing an antisense *tbdn-1* cDNA fragment, which was designed to block endogenous tbdn-1 expression, showed a significant decrease or complete absence of the 69-kDa band representing tbdn-1 (Fig. 1, right). These results indicate that the Ab1272 antibody is specific for detecting tbdn-1 protein in RF/6A cells. The marked decrease of the tbdn-1 band in RF/6A cells harboring *tbdn-1* antisense cDNA is similar to what we had previously shown using IEM cells harboring antisense *tbdn-1* cDNA.[13] These results indicate that tbdn-1 can be specifically detected by Ab1272 in primate choroid-retina endothelial cells as well as in mouse and human endothelial cells.

Tbdn-1 Expression in Ocular Endothelial Homeostasis in Adults

Tbdn-1 immunolocalization was performed in normal adult human eye specimens to determine the levels of tbdn-1 expression in normal adult ocular blood vessels. In four of four normal human adult eye specimens, both limbic (Fig. 2A) and retinal (Fig. 2C, 2E) blood vessels showed high levels of tbdn-1 expression in the endothelial cells lining these vessels. We also detected a very similar pattern of tbdn-1 expression in normal choroidal blood vessel endothelium (see choroidal vessels stained in Fig. 2G). The limbic and retinal blood vessels in normal adult human specimens showed the same staining pattern using an anti-von Willebrand factor antibody (retinal vessels are shown in Figs. 2B, 2D), whereas adjacent sections incubated with either normal rabbit serum or preimmune IgY control antibodies showed no staining (an IgY reacted section is shown in Fig. 2F). These results indicate that, in contrast to most vascular beds, tbdn-1 is expressed at high levels in endothelial linings of normal adult ocular blood vessels during homeostasis.

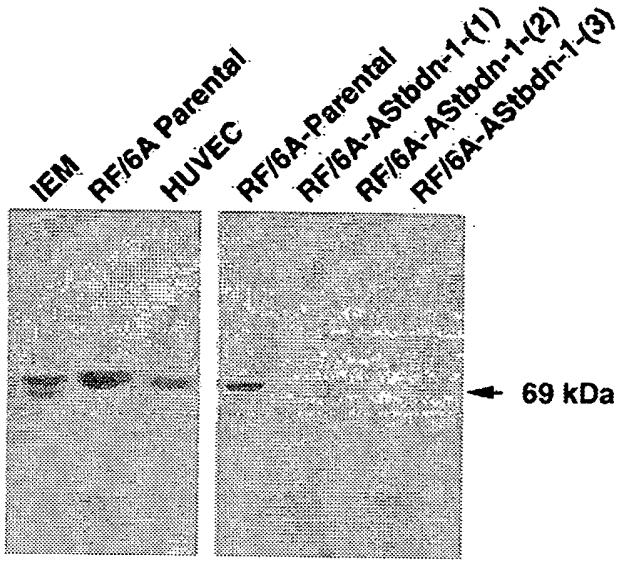


FIGURE 1. Tbdn-1 was specifically detected by anti-tbdn-1 Ab1272 antibody in mouse and human vascular endothelial cells and in rhesus macaque choroid-retina endothelial cells. Expression analysis of tbdn-1 protein in 50 μg of whole-cell lysate of the mouse IEM embryonic vascular endothelial cell line (IEM), the untrasfected rhesus RF/6A choroid-retina endothelial cell line (RF/6A Parental), human umbilical vein endothelial cells (HUVECs), and three separate clones of RF/6A cells stably expressing a *tbdn-1* antisense cDNA fragment (RF/6A-AStbdn-1-3), as indicated. *Arrow*: 69-kDa tbdn-1 band, which resolves as a doublet in the IEM cells.[12]

To assess whether tbdn-1 is expressed by retinal pericytes in vivo, we also analyzed normal human eye sections double stained for tbdn-1 and ASMA, a marker expressed by pericytes and perivascular contractile cells and not by endothelial cells. Figure 3 shows a representative view of a normal human retinal blood vessel double stained for tbdn-1 (Fig. 3, dark brown stain) and ASMA (bright red stain). The tbdn-1 and ASMA stains did not colocalize in retinal blood vessels in normal human eye sections. These results indicate that tbdn-1 does not appear to be expressed in retinal pericytes at the same high levels at which it is expressed in retinal endothelial cells in vivo.

Suppression of Tbdn-1 Expression during Capillary Formation of a Choroid-Retina Endothelial Cell Line

Our previous work has shown that tbdn-1 protein expression is downregulated during capillary formation of the IEM embryonic vascular endothelial cell line in vitro.[13] Because tbdn-1 expression is maintained at high levels in adult ocular blood vessels, contrary to most other vascular beds, we tested in the current study whether tbdn-1 was regulated in a manner different from IEM cells using a model of choroid-retina endothelial cell capillary outgrowth in vitro. We have previously developed an in vitro capillary formation assay using the IEM cell line[13,16,17] and used the RF/6A endothelial cell line derived from rhesus choroid-retina tissue for a similar application in the current study. By treating either IEM or RF/6A cells with an-

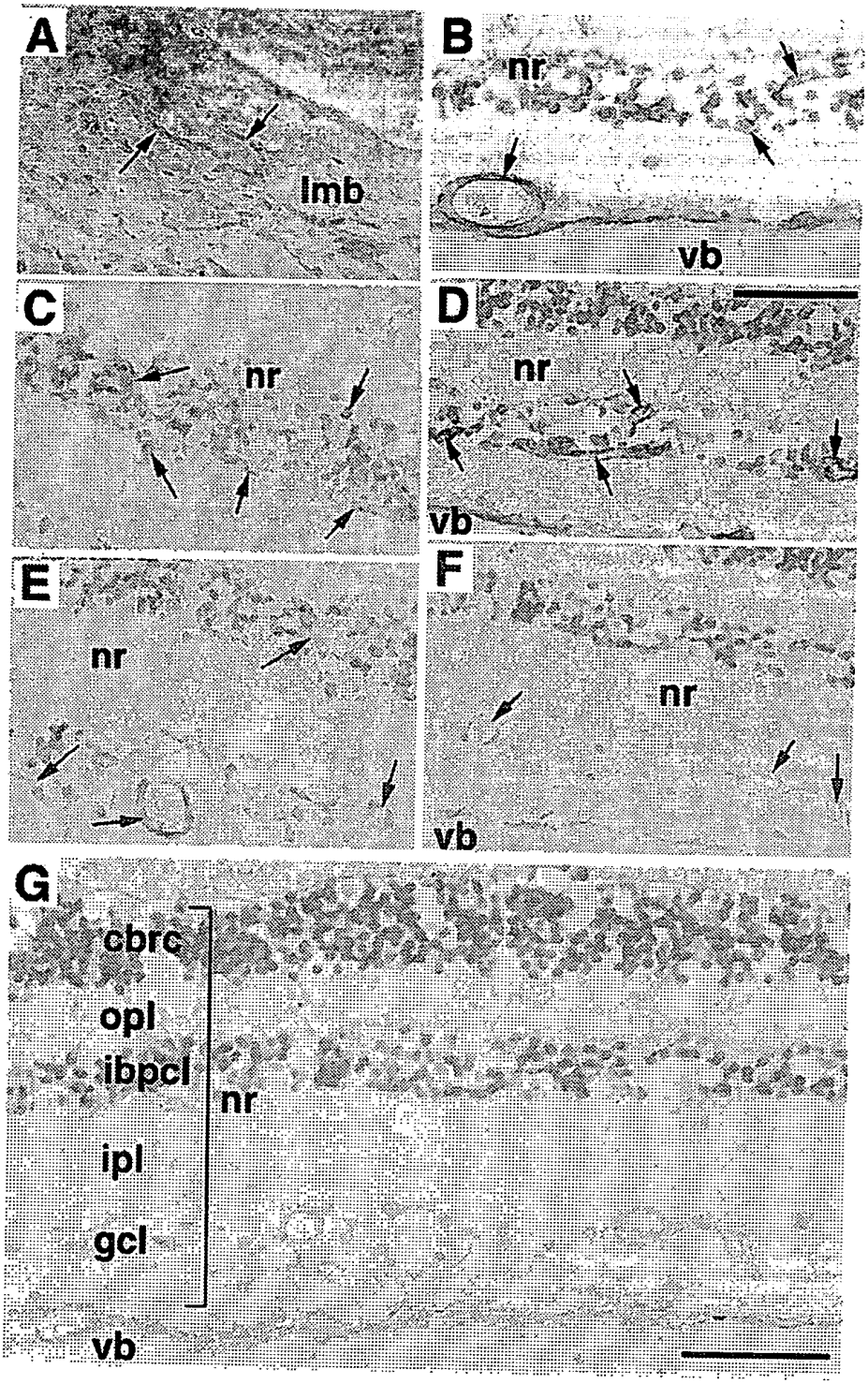


FIGURE 2. Tbdn-1 protein and endothelial marker expression in sections of normal adult human eye. (A) Limbic vessel tbdn-1 expression (*red* stain; *arrows*: tbdn-1–positive endothelial cells in a limbic blood vessel). (C, E) Retinal endothelial tbdn-1 expression in longitudinal- and transverse-sectioned blood vessels in normal adult eye (*red* stain; *arrows*: tbdn-1–positive endothelial cells in retinal blood vessels). (B, D) Retinal endothelial von Willebrand factor expression in longitudinal- and transverse-sectioned blood vessels in normal adult eye (*red* stain, *arrows*: von Willebrand factor–positive endothelial cells in retinal blood vessels). Adjacent sections stained with equal concentrations of preimmune IgY control antibody showed no staining (F). (G) A low-power and labeled view of a methyl green–stained section of the retinal areas shown in (A–F) is provided for orientation purposes. Sections were developed using alkaline phosphatase and fast red substrate; methyl green counterstain. lmb, limbic region of cornea; nr, neural retina; vb, vitreous body; cbrc, cell bodies of rods and cones; opl, outer plexiform layer; ibpcl, integrating bipolar cell layer; ipl, inner plexiform layer; gcl, ganglion cell layer. Scale bar, 50 μm.

giogenic growth factors and then plating the stimulated cells onto a layer of synthetic basement membrane (Matrigel; Collaborative Research, Inc.), we can reproducibly generate colonies of cells sprouting capillary structures (Fig. 4A). These capillary colonies can then be fixed, embedded, and histologically sectioned for immunocytochemical studies as we have previously described for IEM capillary colonies.[13] Of note, little to no staining for tbdn-1 protein was detected in histologic sections of fixed, paraffin-embedded RF/6A cultures induced to form capillary outgrowths in the membrane (Fig. 4 B). However, high levels of tbdn-1 expression were present in histologic sections of fixed, paraffin-embedded RF/6A cells maintained in tissue culture in the absence of any treatment for 48 hours (Fig. 4C). Detection of α-tubulin immunostaining in sections of the preparations of RF/6A capillary cultures (Fig. 4B, inset) confirmed retention of antigenicity in these fixed, paraffin-embedded, and sectioned capillary colony preparations. These results show that a suppression of tbdn-1 expres-

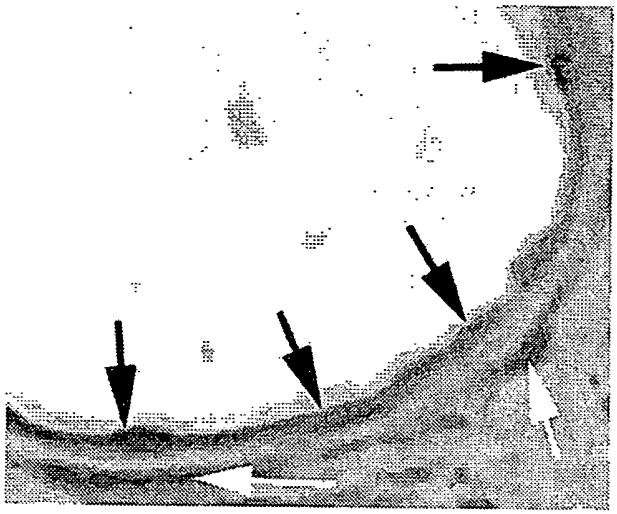


FIGURE 3. Double staining for tbdn-1 and (ASMA) in a retinal vessel of a normal human eye section. Shown is a representative view of a normal human retinal blood vessel double stained for tbdn-1 (*dark brown* peroxidase stain) and ASMA (*bright red* alkaline phosphatase stain). The tbdn-1 and ASMA stains did not colocalize in these retinal blood vessels in normal human eye sections. *Black arrows*: locations of tbdn-1 expression (*brown* staining) in endothelial cells; *white arrows*: locations of ASMA expression (*bright red* staining) in pericyte and perivascular contractile cells. Note that anti-tbdn-1/anti-ASMA double-labeled reactions were overdeveloped to enable clearly revealing the locations of both epitopes.

sion accompanies the induction of capillary formation of RF/6A choroid-retina endothelial cells, in a manner similar to that observed during capillary outgrowth of the IEM cells.

Suppression of Tbdn-1 Expression during Retinal Neovascularization in PDR

Tbdn-1 immunolocalization was performed in diabetic adult human eye specimens in parallel with the normal samples to determine whether the expression characteristics of tbdn-1 in retinal blood vessels change during PDR. Sections of five of five eyes from patients with PDR that were processed and stained simultaneously with the normal human eye samples showed a dramatically lower level of expression of endothelial tbdn-1 protein in the diseased areas of the retinas showing neovascularization. Tbdn-1 was suppressed or completely absent from abnormal proliferating blood vessels and fronds in both preretinal membranes and neural retinal areas in the PDR specimens (Fig. 5C–F for several representative specimens). Quantitative analysis of the red chromogen representing tbdn-1 staining in normal versus PDR specimens by computer (Photoshop; Adobe)[18] revealed a significant decrease of tbdn-1 staining in PDR (10,663 ± 4,740 mean pixels per high-power field; $n = 5$ separate fields from five different eyes) versus normal eyes (173,325 ± 31,042 mean pixels per high-power field; $n = 5$ separate fields from four different eyes). The difference in tbdn-1 staining between normal and PDR specimens was significant ($P < 0.01$ by Student's t-test). PDR specimens showed no change in tbdn-1 levels in the limbic vessels in the anterior portion of the eye in the same sections (compare Fig. 5C–F and 4B). Thus, the suppression of tbdn-1 expression occurred in blood vessels within the neural retina and preretinal membranes but did not occur in limbic vessels in the anterior portions of the same PDR specimens. The limbic vessel expression of tbdn-1 in PDR also served as an internal positive control for tbdn-1 expression in these specimens. We also observed that tbdn-1 was downregulated in the choroidal vessels in the PDR specimens in comparison to choroidal vessels in normal specimens (Fig. 5, low-power view). Expression of the endothelial marker von Willebrand factor was detected at high levels, similar to normal retinal blood vessels, in blood vessels showing decreased tbdn-1 expression from the same PDR specimens (Fig. 5C, 5F, insets). These results indicate that tbdn-1 expression is suppressed in abnormal proliferating blood vessels of the neural retina and preretinal membranes in PDR.

DISCUSSION

Tbdn-1 expression peaks during early to middle stages of development of most blood vessels and is downregulated at later stages of maturation, suggesting it may be involved with regulating specific stages of blood vessel maturation during embryogenesis.[13] This is exemplified by tbdn-1 expression in yolk sac vasculature development, in which tbdn-1 is expressed most highly during early stages of yolk sac vasculature formation and is downregulated at the later stages of development during which time angiogenesis of the vitelline vasculature occurs.[13]

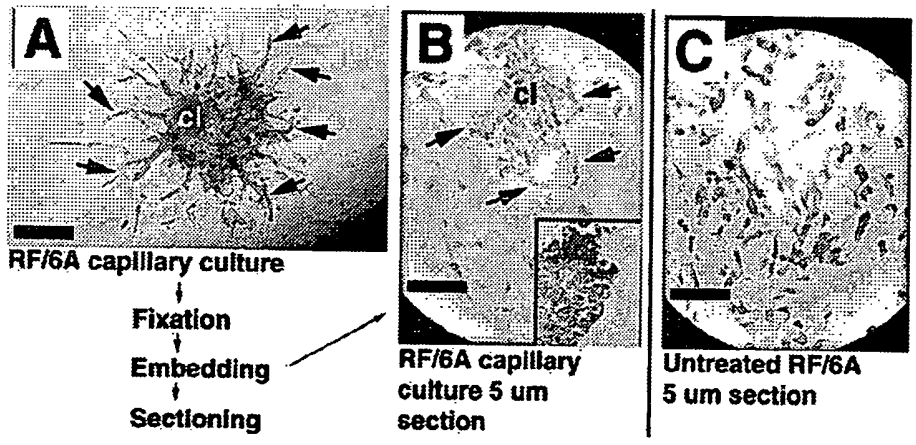


FIGURE 4. Suppression of tbdn-1 protein expression in RF/6A cells in vitro during induction of capillary formation on synthetic basement membrane. (A) RF/6A capillary colonies shown in culture before processing. (B) After fixation, embedding, sectioning, and staining with anti-tbdn-1 Ab1272, RF/6A capillary colonies showed low staining levels for tbdn-1. (B, *arrows*) Similar capillary sprouts as indicated by *arrows* in (A); cl indicates main body of the colony. Sections of RF/6A capillary colonies from the same preparation stained very strongly with anti-tubulin positive control antibody (B, *inset, dark red* stain). (C) Sections of RF/6A cells harvested from untreated cultures that were fixed and paraffin embedded were highly positive when stained using anti-tbdn-1 Ab1272 (*dark red* stain). Sections of RF/6A cells and capillary colonies stained with equal concentrations of preimmune IgY were negative (not shown). Staining of sections was developed using alkaline phosphatase and fast red substrate. Methyl green counterstain in (B) reveals the capillary sprouts (shown by *arrows* in the capillary colony before processing, in A, and after processing, in B). cl, sprouting from the RF/6A colonies. Scale bar, 50 μm.

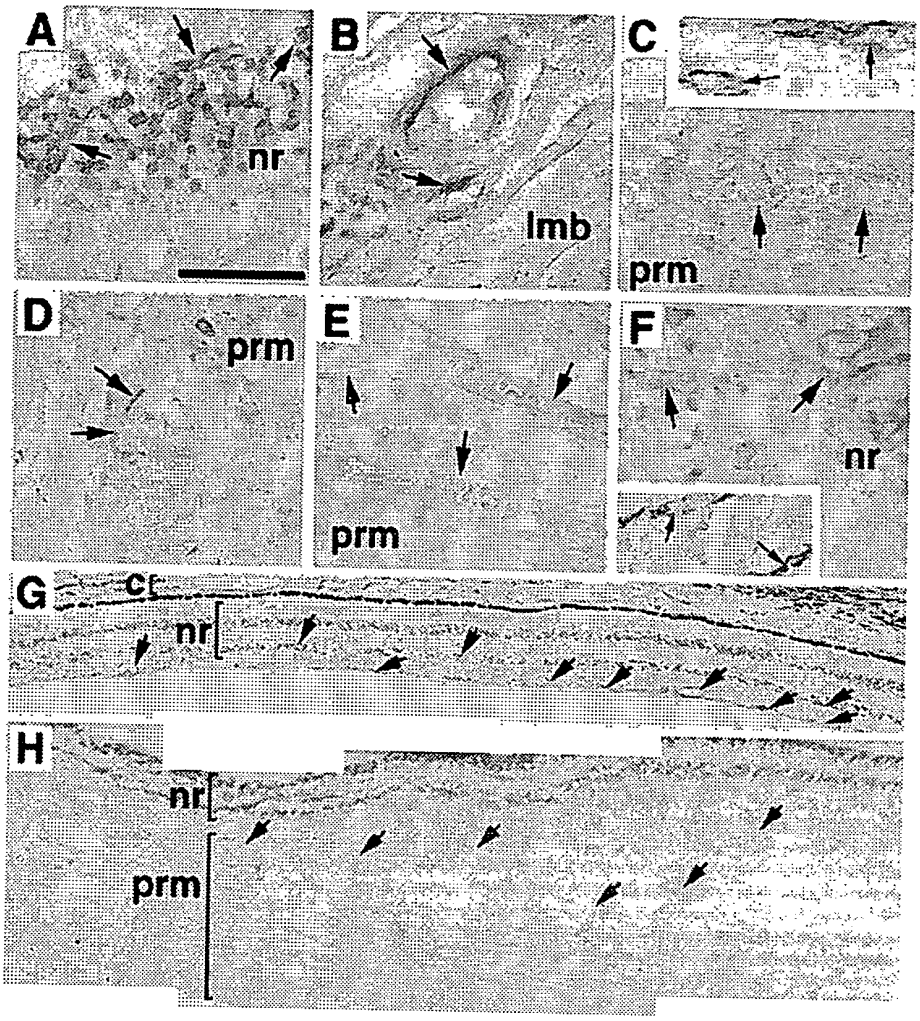


FIGURE 5. Tbdn-1 protein expression was suppressed in specimens of eyes from patients with PDR. (A) Retinal endothelial tbdn-1 expression (*arrows*: retinal blood vessels stained *red*) in normal adult eye. (C–E) Tbdn-1 staining in blood vessels in preretinal membranes in sections of eyes from three separate representative patients with PDR. (F) Tbdn-1 staining in blood vessel fronds cut longitudinally in a neural retinal area in a section of eye from a fourth and separate representative patient. (C, F, *insets*) von Willebrand factor staining of abnormal blood vessels (*arrows*) in sections from the same PDR specimens and adjacent to those stained for tbdn-1. Blood vessels in the diseased retinal tissue showed either very low levels of tbdn-1 expression or no detectable tbdn-1 expression, compared with normal specimens, whereas the same abnormal blood vessels expressed von Willebrand factor (see also the Results section for quantitative analysis of tbdn-1 expression levels in these sections). (B) Tbdn-1 staining (*arrow*, *red*) of limbic blood vessels in the anterior part of the same section as that shown in (D) to exemplify normal tbdn-1 expression in unaffected areas of PDR-affected eyes. All sections were also incubated with equal concentrations of preimmune IgY and showed no staining (see example in Fig. 2). Low-power views of a normal retina (G) and a diabetic retina with a preretinal membrane (H), both stained for tbdn-1 are provided for orientation purposes. Sections were developed using alkaline phosphatase and fast red substrate with methyl green counterstain. lmb, limbic region of cornea; nr, neural retina; vb, vitreous body; c, choriocapillaris; preretinal membrane. Scale bar, 50 μm.

In the adult, tbdn-1 is not ubiquitously expressed in all blood vessels but is restricted to the endothelium of highly specialized vascular beds (e.g., atretic ovarian vasculature and atrial endocardium.[13]).

These studies suggest that tbdn-1 may play a role in some specialized vascular beds during adulthood as well. The results of the present study provides two lines of evidence to suggest that tbdn-1 expression may be involved in maintaining ocular blood vessel homeostasis. First, tbdn-1 expression persisted at high levels in normal adult ocular blood vessels. Second, retinal endothelial tbdn-1 expression was suppressed in neovascularization of PDR. In interpreting these data, it could be argued that the observed loss of tbdn-1 was due simply to loss of endothelial cells or, alternatively, to loss of pericytes, if it were the case that tbdn-1 was also expressed by retinal pericytes. Our present results and the results of others[19] have shown that diseased blood vessels in PDR specimens retain expression of the endothelial marker von Willebrand factor, indicating that the decrease in tbdn-1 expression is not merely a consequence of loss of all vascular endothelial cells in these blood vessels, as has been reported for some vascular beds in certain stages of retinopathy in other studies.[20,21]

Furthermore, to assess whether tbdn-1 could be expressed by pericytes in vivo, we analyzed human eye sections double stained for tbdn-1 and ASMA, a cytoskeletal isoform of vascular actin expressed by pericytes and nonendothelial perivascular contractile cells.[22] Because these markers did not show an obvious colocalization pattern in normal human retina sections, retinal pericytes do not appear to express the same high levels of tbdn-1 as found in retinal endothelial cells in vivo. Our results do not exclude the possibility that tbdn-1 may be expressed in pericytes at very low levels below the limit of detection by these methods.

Nevertheless, all these results taken together indicate that suppression of retinal blood vessel tbdn-1 expression in neovascularization of PDR is a reflection of a decrease in tbdn-1 levels in retinal endothelial cells rather than a reflection of cell loss. Because tbdn-1 is expressed in normal retinal endothelium but is suppressed in retinal endothelium of PDR, our results prompt speculation that a possible functional role for tbdn-1 in normal retinal capillaries may be to participate in a mechanism that may dampen capillary outgrowth. Conversely, because tbdn-1 suppression is associated with the abnormal retinal capillary outgrowth occurring during neovascularization in PDR, removal of such a potential dampening influence of tbdn-1 may permit outgrowth of retinal capillaries in the diabetic environment.

To study tbdn-1 in ocular endothelium in vitro, we used the rhesus RF/6A choroid-retina endothelial cell line[14,15] as a model system. The fact that the primate RF/6A cell line is evolutionarily closer to human than the mouse or bovine renders the RF/6A model more attractive than nonprimate models. Furthermore, RF/6A cells were derived by spontaneous immortalization rather than with the use of exogenous transforming oncogenes. However, it is not known whether RF/6A cells were derived entirely from the choroid, the retina, or a mixture of both structures.[14,15] Therefore, although RF/6A cells may not be a genuine representation of either choroid or retina endothelial cells as they occur in vivo, this may not be a critical factor for the purpose of our study, because tbdn-1 is expressed in both retinal and choroidal vessels in vivo. The RF/6A cell line possesses a number of properties consistent with and characteristic of vascular endothelium,[14,15] and our unpublished observation, 2000, and thus retains some value as an in vitro model system for studies of tbdn-1. Tbdn-1 expression during in vitro capillary outgrowth of RF/6A cells was significantly reduced or absent compared with control cells growing under normal, unstimulated culture conditions in vitro. Our results obtained with the choroid-retina RF/6A in vitro model correlate with downregulation of tbdn-1 during IEM cell in vitro capillary outgrowth.[13]

Most important, downregulation of tbdn-1 during capillary outgrowth of RF/6A correlates with the suppression of tbdn-1 expression we observed in abnormal neural retinal blood vessels, blood vessels and fronds in preretinal membranes, and choroidal blood vessels in PDR specimens. Although retinal and choroidal capillaries are anatomically and physiologically different, choroidal pathologic neovascularization occurs in PDR.[20] The suppression of tbdn-1 in diseased retinal and choroidal vessels is consistent with the pathologic course of PDR. Because the regulation pattern of tbdn-1 in RF/6A cell capillary outgrowth correlates with the regulation pattern of tbdn-1 in retinal neovascularization of PDR, our results indicate that the RF/6A system may serve as a useful model for studies of retinal capillary outgrowth.

Our results suggest that the microenvironment in the disease-affected regions in PDR retinas may harbor a local milieu that supports the downregulation of tbdn-1, in that limbic vascular tbdn-1 levels were not different from normal in the PDR specimens we analyzed (see the Results section and Fig. 5, comparing 5C–F with 5A and 5B). This hypothesis is also supported by our observation that both blood vessels and capillary fronds showed a suppression of tbdn-1 expression in the tissues of PDR-affected retinas. The PDR microenvironment may include factors present in PDR retinal tissue that may lead to downregulation of tbdn-1 levels. Furthermore, the abnormal death of cells such as pericytes in the retinal vascular wall may cause derangements in the diabetic retinal microenvironment to which the remaining and viable retinal endothelial cells become exposed.[21]

A range of angiogenic growth factors (VEGF, bFGF, and IGF-1), integrins, and derangements of extracellular matrix (ECM) components (such as collagen type IV) are associated with pathologic neovascularization in PDR, any or all of which could potentially affect tbdn-1 expression.[3–5,23–27] However, our recent data suggest that tbdn-1 expression levels may be altered by ECM components rather than by direct actions of angiogenic growth factors, such as VEGF and bFGF (Paradis H, Gendron RL, unpublished observations, 2001). Our observation of a similar suppression of tbdn-1 expression during RF/6A choroid-retina capillary outgrowth in vitro is consistent with this evidence, because the synthetic membrane used is a reconstitution of basement membrane components (Matrigel; Collaborative Research, Inc.) and is known to contain a range of ECM components, such as collagen type IV, heparan sulfate proteoglycans, laminin, and entactin.[28,29]

Despite the likely caveats associated with interpreting the regulation of endothelial behavior in reconstitution experiments in vitro and during PDR in vivo, our results indicate a correlation between suppression of tbdn-1 expression and retinal capillary formation occurring in choroid-retina capillary outgrowth in vitro and during neovascularization of PDR in vivo. We are currently in the process of identifying the ECM components that may regulate tbdn-1 expression.

Of particular interest, the expression of tbdn-1 in normal adult retinal blood vessels parallels the expression of pigment epithelium derived factor (PEDF) in adult retina, a recently described novel antiangiogenic serpin family member produced by the normal retinal pigment epithelium. Decreases in the expression levels of PEDF have been observed during oxygen-induced retinal neovascularization in mice and rats,[30,31] and systemic administration of PEDF to mice with ischemia-induced retinopathy prevents retinal neovascularization in this model.[32] It has not yet been determined whether PEDF expression levels are decreased in retinal tissues in human PDR specimens, but it can be predicted that this would be the case. We also do not know at this time whether tbdn-1 can be regulated either directly or indirectly by PEDF.

Although animal models of retinal neovascularization have been studied, little information is available about the intracellular mechanisms in retinal vascular cells that are associated with neovascularization during PDR in human specimens. Polymorphisms of the aldose reductase gene, which may alter aldose reductase mRNA levels within cells, are thought to predispose patients with diabetes to retinopathy through possible disturbances in the polyol pathway and subsequent vascular damage.[33] In diabetes, increases in retinal pericyte expression levels of the proapoptotic protein Bax have been associated in a recent study with increased apoptotic death of retinal pericyte cells.[34] Although they detected no changes in endothelial Bax expression in the specimens analyzed, the investigators indicate that Bax may be regulated with a different time course in retinal endothelial cells in diabetes.[34]

Our finding of high levels of tbdn-1 expression in adult ocular blood vessel endothelial cells during homeostasis and the loss of this expression of tbdn-1 during retinal capillary outgrowth occurring in PDR sheds light on the intracellular processes that are disregulated during neovascularization associated with PDR. The re-expression of tbdn-1 in diseased vessels in PDR may be necessary to restore homeostasis and stop neovascularization. Tbdn-1 is associated with an acetyltransferase activity and contains protein–protein interaction and DNA binding-like motifs.[13] Therefore, if tbdn-1 is indeed mechanistically involved in regulating neovascularization in the eye, it can be speculated that it may act through acetyltransferase activity and/or protein–protein interactions similar to its yeast homologue, the N-terminal acetyltransferase subunit NAT1.[35–38]

Acknowledgments

The authors thank Candace Kao for expert technical assistance and Dale L. Phelps, Winston Kao, and Chia-Yang Liu for helpful discussions.

References

1. Kohner EM. Diabetic retinopathy. *BMJ.* 1993;307:1195–1199.
2. Infeld DA, O'Shea JG. Diabetic retinopathy. *Postgrad Med J.* 1998; 74:129–133.
3. Paques M, Massin P, Gaudric A. Growth factors and diabetic retinopathy. *Diabetes Metab.* 1997;23:125–30.
4. Miller JW. Vascular endothelial growth factor and ocular neovascularization. *Am J Pathol.* 1997;151:13–23.

5. Robinson GS, Aiello LP. Angiogenic factors in diabetic ocular disease: mechanisms of today, therapies for tomorrow. *Int Ophthalmol Clin.* 1998;38:89-102.
6. Hammes HP, Lin J, Bretze RG, Brownlee M, Breier G. Upregulation of the vascular endothelial growth factor/vascular endothelial growth factor receptor system in experimental background diabetic retinopathy of the rat. *Diabetes.* 1998;47:401-406.
7. Amin RH, Frank RN, Kennedy A, Eliott D, Puklin JE, Abrams GW. Vascular endothelial growth factor is present in glial cells of the retina and optic nerve of human subjects with nonproliferative diabetic retinopathy. *Invest Ophthalmol Vis Sci.* 1997;38:36-47.
8. Ozaki H, Seo MS, Ozaki K, et al. Blockade of vascular endothelial cell growth factor receptor signaling is sufficient to completely prevent retinal neovascularization. *Am J Pathol.* 2000;156:697-707.
9. Robinson GS, Pierce EA, Rook SL, Foley E, Webb R, Smith LES. Oligodeoxynucleotides inhibit retinal neovascularization in a murine model of proliferative retinopathy. *Proc Natl Acad Sci USA.* 1996;93:4851-4856.
10. Aiello LP, Pierce EA, Foley ED, et al. Suppression of retinal neovascularization in vivo by inhibition of vascular endothelial growth factor (VEGF) using soluble VEGF-receptor chimeric proteins. *Proc Natl Acad Sci USA.* 1995;92:10457-10461.
11. Luna J, Tobe T, Mousa SA, Reilly TM, Campochiaro PA. Antagonists of integrin a-v b-3 inhibit retinal neovascularization in a murine model. *Lab Invest.* 1996;75:563-573.
12. Hammes H, Brownlee M, Jonczyk A, Sutter A, Preissner K. Subcutaneous injection of a cyclic peptide antagonist of vitronectin receptor-type integrins inhibits retinal neovascularization. *Nat Med.* 1996;2:529-533.
13. Gendron RL, Adams LC, Paradis H. *Tubedown-1,* a novel acetyltransferase associated with blood vessel development. *Dev Dyn.* 2000;218:300-315.
14. Lou DA, Hu FN. Co-distribution of von Willebrand factor and fibronectin in cultured rhesus endothelial cells. *Histochem J.* 1987;19:431-438.
15. Lou DA, Hu FN. Specific antigen and organelle expression of a long-term rhesus endothelial cell line. *In Vitro Cell Dev Biol.* 1987;23:75-85.
16. Gendron RL, Tsai, F-Y, Paradis H, Arceci RJ. Induction of embryonic vasculogenesis by bFGF and LIF in vitro and in vivo. *Dev Biol.* 1996;177:332-347.
17. Paradis H, Gendron RL. LIF transduces contradictory signals on capillary outgrowth through induction of Stat3 and $p^{41/43}$ MAP kinase. *J Cell Sci.* 2000;113:4331-4339.
18. Lehr HA, van der Loos CM, Teeling P, Gown AM. Complete chromogen separation and analysis in double immunohistochemical stains using Photoshop-based image analysis. *J Histochem Cytochem.* 1999;47:119-126.
19. Boeri D, Cagliero E, Podesta F, Lorenzi M. Vascular wall von Willebrand factor in human diabetic retinopathy. *Invest Ophthalmol Vis Sci.* 1994;35:600-607.
20. Lutty GA, Cao J, McLeod DS. Relationship of polymorphonuclear leukocytes to capillary dropout in the human diabetic choroid. *Am J Pathol.* 1997;151:707-714.
21. Mizutani M, Kern TS, Lorenzi M. Accelerated death of retinal microvascular cells in human and experimental diabetic retinopathy. *J Clin Invest.* 1996;97:2883-2890.
22. Sawtell NM, Lessard JL. Cellular distribution of smooth muscle actins during mammalian embryogenesis: expression of the alpha-vascular but not the gamma-enteric isoform in differentiating striated myocytes. *J Cell Biol.* 1989;109:2929-2937.
23. Frank RN, Amin RH, Eliott D, Puklin JE, Abrams GW. Basic fibroblast growth factor and vascular endothelial growth factor are present in epiretinal and choroidal neovascular membranes. *Am J Ophthalmol.* 1996;122:393-403.
24. Kon CH, Occleston NL, Aylward GW, Khaw PT. Expression of vitreous cytokines in proliferative vitreoretinopathy: a prospective study. *Invest Ophthalmol Vis Sci.* 1999;40:705-712.
25. Sivalingam A, Kenney J, Brown GC, Benson WE, Donoso L. Basic fibroblast growth factor levels in the vitreous of patients with proliferative diabetic retinopathy. *Arch Ophthalmol.* 1990;108:869-872.
26. Boulton M, Gregor Z, McLeod D, et al. Intravitreal growth factors in proliferative diabetic retinopathy: correlation with neovascular activity and glycaemic management. *Br J Ophthalmol.* 1997;81:228-233.
27. Ljubimov AV, Burgeson RE, Butkowski RJ, et al. Basement membrane abnormalities in human eyes with diabetic retinopathy. *J Histochem Cytochem.* 1996;44:1469-1479.
28. Vukicevic S, Kleinman HK, Luyten FP, Roberts AB, Roche NS, Reddi AH. Identification of multiple active growth factors in basement membrane Matrigel suggests caution in interpretation of cellular activity related to extracellular matrix components. *Exp Cell Res.* 1992;202:1-8.
29. Kleinman HK, McGarvey ML, Liotta LA, Robey PG, Tryggvason K, Martin GR. Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan from the EHS sarcoma. *Biochemistry.* 1982;21:6188-6193.
30. Dawson DW, Volpert OV, Gillis P, et al. Pigment epithelium-derived factor: a potent inhibitor of angiogenesis. *Science.* 1999;285:245-248.
31. Gao G, Li Y, Zhang D, Gee S, Crosson C, Ma J. Unbalanced expression of VEGF and PEDF in ischemia-induced retinal neovascularization. *FEBS Lett.* 2001;489:270-276.
32. Stellmach VV, Crawford SE, Zhou W, Bouck N. Prevention of ischemia-induced retinopathy by the natural ocular antiangiogenic agent pigment epithelium-derived factor. *Proc Natl Acad Sci USA.* 2001;98:2593-2597.
33. Demaine A, Cross D, Millward A. Polymorphisms of the aldose reductase gene and susceptibility to retinopathy in type 1 diabetes mellitus. *Invest Ophthalmol Vis Sci.* 2000;41:4064-4068.
34. Podesta F, Romeo G, Liu WH, et al. Bax is increased in the retina of diabetic subjects and is associated with pericyte apoptosis in vivo and in vitro. *Am J Pathol.* 2000;156:1025-1032.
35. Mullen JR, Kayne PS, Moerschell RP, et al. Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. *EMBO J.* 1989;8:2067-2075.
36. Whiteway M, Freedman R, Van Arsdell S, Szostak JW, Thorner J. The yeast ARD1 gene product is required for repression of cryptic mating-type information at the HML locus. *Mol Cell Biol.* 1987;7:3713-3722.
37. Park EC, Szostak JW. ARD1 and NAT1 proteins form a complex that has N terminal acetyltransferase activity. *EMBO J.* 1992;11:2087-2093.
38. Aparicio OM, Billington BL, Gottschling DE. Modifiers of position effect are shared between telomeric and silent mating-type loci in S. cerevisiae. *Cell* 1991;66:1279-1287.